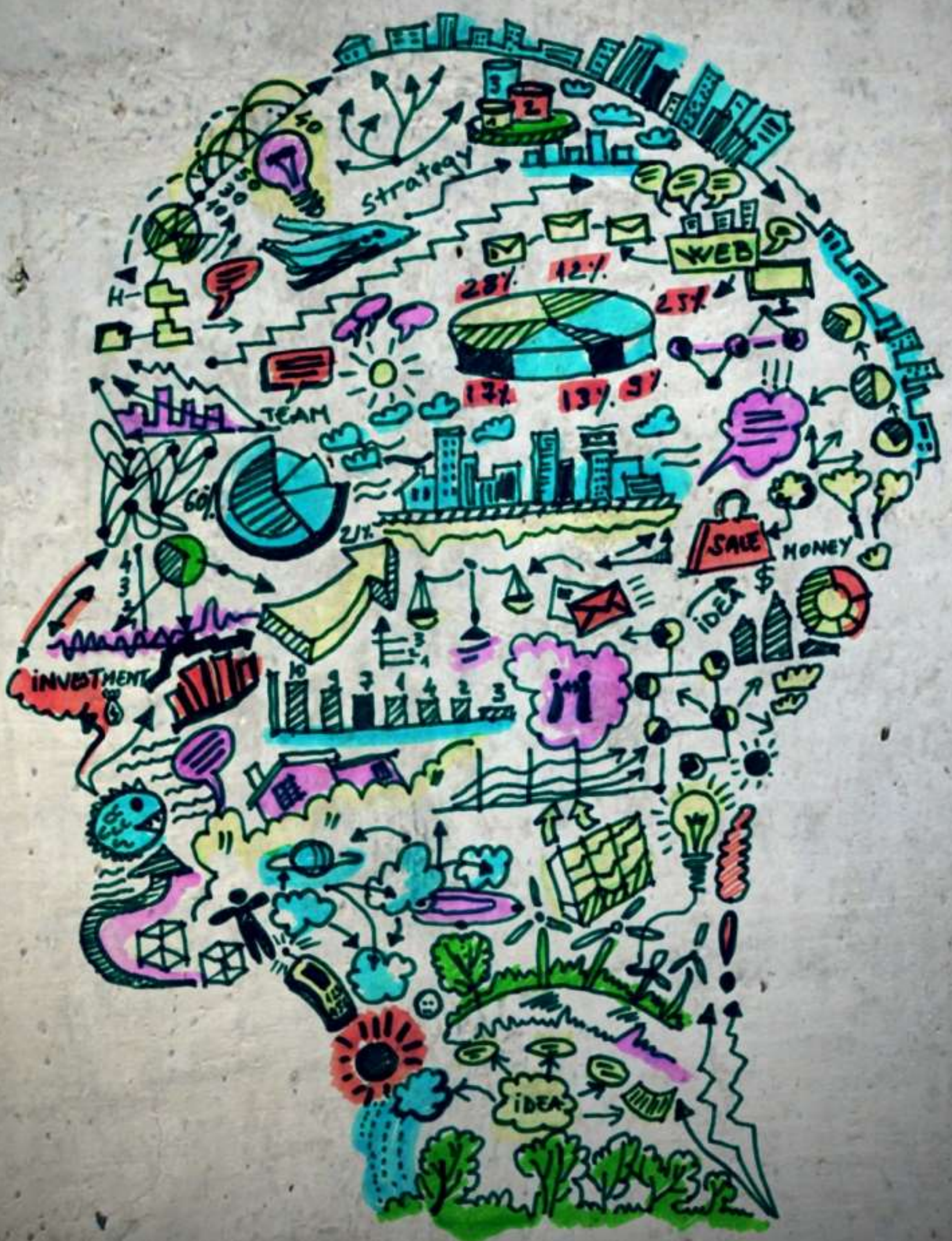
Ipsos MORI
Social Research Institute



March 2016

The One-Way Mirror: Public attitudes to commercial access to health data

Report prepared for the Wellcome Trust

wellcome trust

Contents

List of Figures

List of Tables

1 Summary of report

1 Summary of Report

1.1 Introduction

Data is collected throughout the health service in increasingly large quantities, as well as in the contexts of biomedical and health research. Data is collected for direct care and for secondary uses as well.

In surveys, such as Ipsos MORI's work for the Royal Statistical Society[1] (RSS) the public say that in principle they do not want their health records being shared with private companies. However, there are many different types of commercial access going on now, and possible in future. The Wellcome Trust therefore wanted to investigate how, and in what ways, the public would distinguish between different types of commercial access; and whether the type of data used, and the types of data user, would have an influence on the level of acceptability to the public.

The Trust therefore commissioned Ipsos MORI to carry out research to understand how attitudes towards commercial access to health data are formed and influenced, among a cross-section of the British general public and with specific audiences such as healthcare professionals, patients, and members of cohort studies.

1.1.1 Objectives

The **primary objectives** of this research were to:

- Identify **factors influencing attitudes** towards commercial organisations accessing health, biomedical and genetic data;

- Identify governance, safeguarding and communications actions that could help improve trustworthiness of research uses and protections of data; and enable public trust in access to data to be developed over time;

- Provide an **evidence base for the Wellcome Trust** to draw upon as it considers how best to develop policy and engage other relevant bodies, research teams and the public on these issues.

Secondary objectives were:

- In order to best understand the detail of the factors influencing attitudes, to explore how people perceive the **relative importance of different variables** in sharing health data with commercial bodies.

---

[1] RSS/Ipsos MORI (2014), 'Public attitudes to data sharing'; https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-lessons-for-policymakers.aspx

- Identify any emerging differences between **sub groups and different audiences**.
- Quantify views on key topics which emerge from the qualitative phase.

1.1.2 Project design

Extensive and robust qualitative social research study. Sixteen qualitative workshops across Great Britain with 246 individuals in total:

- Eight full-day deliberative workshops with members of the **general public**;
- Three evening workshops with **GPs and hospital doctors**;
- Four evening/daytime workshops with people who have a **long-term condition**, including rare diseases;
- One full-day workshop with 'Children of the 90s' cohort members.

During the qualitative workshops, participants were introduced to **six case studies** (or four, in the shorter workshops). The case studies gave examples of existing data-sharing activity in the public and private sectors, illustrated different variables that might make a difference and set out the full range of types of commercial access to data. They were followed up with discussion of hypothetical future activities (including some deliberately provocative and controversial ones). These examples were used as springboards for discussion.

A **digital anthropologist** attended one workshop to provide an analysis of the conversation from this perspective. **Improvisational actors** attended two workshops to bring to life in comedy sketches some of the most emotive issues as they arose.

The qualitative work was followed up by a **quantitative survey** that examined some of the issues arising from the workshops. The questionnaire was designed in conjunction with the Wellcome Trust and the advisory group, and aimed to fill in gaps that previous surveys have left open on the subject of data sharing with commercial organisations.

The survey focussed on topics about acceptability with sharing of health data in different situations and included questions on:

- **Awareness** of different organisations using health data;
- **Support** for sharing health data with commercial organisations for research purposes;
- How support for this sharing differs depending on the **specifics of the organisation**;
- Different **safeguards** and factors that could impact views of acceptability; and
- Attitudes towards **consent and permission** in the data sharing process.

Ipsos MORI interviewed a representative quota sample of 2,017 adults across Great Britain aged 16 and over. Face-to-face interviews were conducted in-home between 30 November and 11 December 2015. Data is weighted to the known population profile. The results are described in chapter six, and topline findings can be found in the appendix.

NB: The qualitative findings do not exactly map and mirror the quantitative ones.

Qual participants went on a much more substantive journey through the day and their views were very nuanced. In the qual, there was more scope for getting information about details while in the quant there was no discussion or context given.

The report links the two, but they stand alone. We have indicated where parallels are striking.

1.2 Context

Public views show low understanding of data and low knowledge of how healthcare works. In the deliberative work, this led to wariness and scepticism about the idea of commercial access to healthcare data.

In the deliberative workshops, participants had limited knowledge of some important contexts. These included:

What is data? Participants made no division made between health and medical data; both were seen as personal and individual. Participants believed that they **owned their personal, individual-level data** both in the arena of health, and more generally. They imagined that the **default option** would be that all such individual-level data ought not to be passed on if shared, that sharing should be consented, and that it should only ever take place if there was a good reason.

"Statistics" was better understood than "aggregate". Participants understood 'statistics' to be the numbers taken from aggregate data, where no individual data lines could be identified. Some socially minded people saw risks in sharing aggregate data, even though it could not be identified, if it was to be used to segment or disadvantage vulnerable groups.

Individual level data is felt to be "My data" even if anonymised. Many saw risks in individual data sharing as they believe it can be de-anonymised. This may have detrimental effects on the individual (for example discrimination or financial loss). However, when data is passively collected, it is less likely to be seen as 'my data'; sharing this kind of data is seen as having little risk to the individual (though for some, it could prove a risk to society).

Healthcare professionals had a more nuanced understanding of the idea of data ownership and permission to use.

Genetic data was not known about or understood, and the potential for its use, broadly, was not known (patients and healthcare professionals had more sense of its potential).

The **survey** also demonstrated that **detailed awareness of how the NHS uses health data is low**, with just a third (33 per cent) reporting to have heard a great deal or a fair amount about how the NHS is using health data. This awareness falls further still to just 16 per cent and 18 per cent awareness for commercial organisations and academics, respectively.

Participants in the workshops did not know that commercial companies already play a part in delivering healthcare and biomedical research. They also did not spontaneously mention academics and charities as part of the health system. They did not draw a meaningful distinction between private companies' research and retail arms.

Some were shocked to hear that private companies were engaged at all with healthcare. The motivations of commercial companies in delivering health services were questioned in deliberation and the private sector in general was mistrusted. Many approached the discussion of data sharing with some caution, as a result.

Some participants felt that private sector involvement suggested a wider agenda of the NHS in general 'being privatised'. They did not want this, so were broadly against the idea of data sharing in principle.

There was little detailed understanding of safeguarding practices. There was no awareness of the current regulatory framework around the use of anonymised health data. Participants showed very low knowledge of safeguarding practices (partly related to their lack of knowledge around how datasets are used and managed).

Currently there is no form of opt-out from anonymised health data usage; but participants did not know this.

Healthcare professionals and patients knew more; they raised issues such as the **quality of data** and how well it is collected.

The **quantitative** work found that several interconnected factors appear related to acceptance of commercial access. **Educational attainment, awareness of data usage and social grade all appear to be linked to acceptance of commercial access**, in many places in the quantitative study.

In the survey findings, **greater knowledge about the subject and exposure to the ideas tends to be related to acceptance.**

The survey also found that the **relationship between acceptance of commercial access to health data and age is non-linear**. While the relationship between age and acceptance is complex, it is clear that **young people are not automatically more in favour of commercial access;** other factors play a more important role in driving acceptance.

The background context led to some **confusion and wariness around data sharing** in the workshops. Overall, the fear could be summed up in the idea that current data sharing seems unidirectional, shadowy and hard to understand.

'It's a *one-way mirror****; they know everything about you, but we don't know what they're doing with that.'***
Patients (severe conditions), London

Contexts around data are also changing. A 'context collapse' is underway

Through analysis of the way people talked in the qualitative sessions, and using the lens of digital anthropology provided by our supporting study, we assert that the qualitative participants were trying to make sense of **changing ways that data is shared**.

In particular, they have a sense that there are **new, uncharted kinds of data transactions**, in new kinds of contexts, today.

There are two traditional contexts of data sharing which people know:

- **Consumer transactions**: **BUYING** online, where data is actively given as part of the commercial deal (e.g. getting online special offers by giving your email address); or **DOING** activities where data is passively collected (e.g. using a smart phone to track your fitness, knowing that the phone company will also repurpose passively given data).

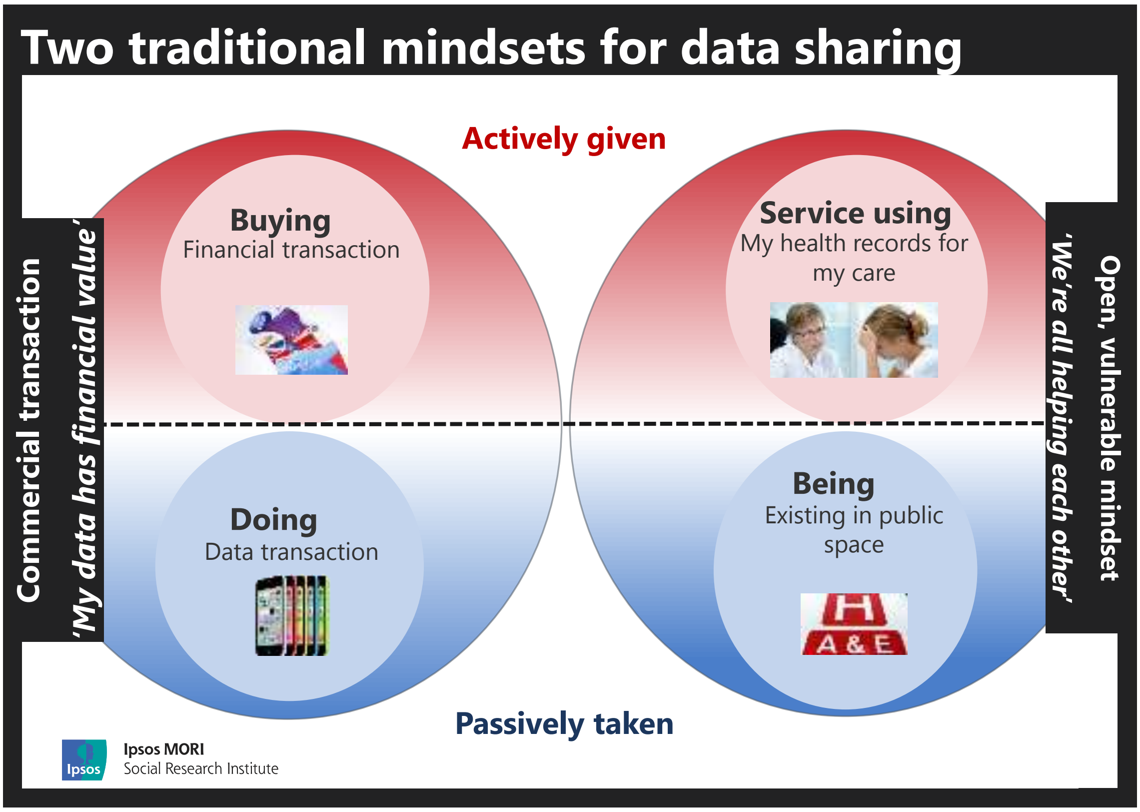
- **Participants believe that the companies will sell on any data they possibly can from these transactions.**

- **Social Contract experiences**: **SERVICE USING** (e.g. going to the doctor, giving information actively for your care); or just **BEING**, participating in (non-commercial) life using a hospital, a road, a library or other public space, generating data **passively**).

- **Participants underestimate the amount of data that is collected in these circumstances and imagine that only very light touch, basic, aggregated statistical analysis usually takes place.**

These two traditional contexts, and the mindsets which typically correspond to them, are illustrated in figure 1.1.

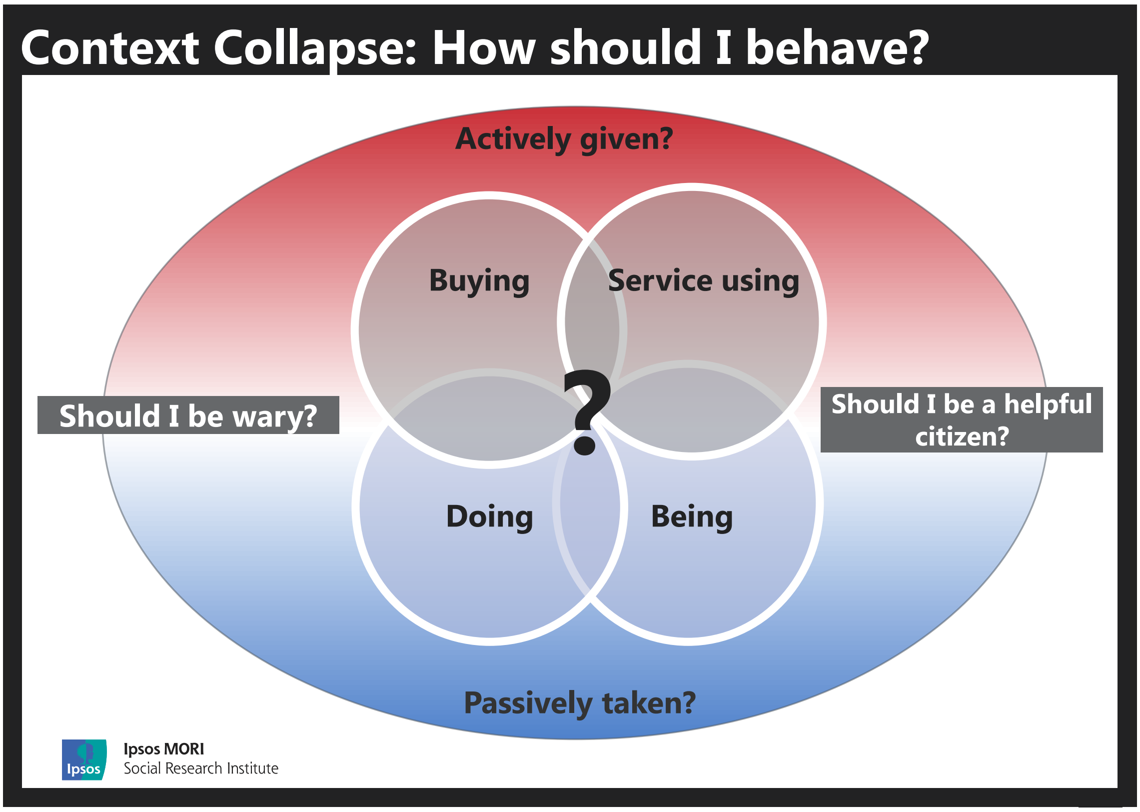
Figure 1.1 – The two traditional contexts/mindsets for data sharing



Uses of health data in commercial contexts cause these two traditional contexts to collapse together. Understanding this 'context collapse' is vital for understanding why the public approves or disapproves of different types of commercial access to health data.

This 'context collapse' is illustrated in figure 1.2.

Figure 1.2 – The 'context collapse'



Participants in the workshops were **confused** about what behaviours and assumptions these new contexts required, and therefore were worried and suspicious that they might be harmed.

The public in deliberations tried to make sense of the changing world by reverting back to assumptions and prejudices (healthcare should be delivered by the NHS, companies are untrustworthy) in order to feel more secure.

Participants expressed their fears by saying that vulnerable groups in society, such as the elderly or less digitally literate, who do not know the 'rules' of the commercial context, may be at risk of harm or exploitation.

In fact, **younger participants may also be at risk**; in the workshops, many participants, even the youngest, did not know that content they posted on social media could form part of their data profile and be purchased or used by a range of organisations; they were not as aware of the current situation as they thought they were.

1.3 Key factors influencing attitudes

We used case studies to explore attitudes to variables. These are summarised below.[2]

Summary of the six case studies

- **Data linking and analysis in the NHS** – An NHS trust asks a healthcare intelligence company to analyse individual-level data on patient journeys, to see if there are different patterns in health outcomes, and predict drivers of service use.
- **Monitoring safety of drugs and medicines** – A public health regulator runs an observational study to look at long term side effects of a blood pressure drug. Primary care data is provided to compare the probability of serious adverse events among those taking the drug, compared to those on other drugs for high blood pressure.
- **Calculating insurance premiums** – Private health insurance companies use hospital data about diagnoses and hospital admissions, and find that those living in deprived areas were more likely to develop certain critical illnesses.
- **Using genetic data in care and research** – Patients consent to having their genome sequenced as part of their clinical care. This is linked to their medical records to aid diagnosis and treatment, and made available for research by academics, scientists and commercial organisations.
- **Pharmacists using Summary Care Records** – The NHS wants all community pharmacists to have access to a summary care record. Pharmacists would have access to this with patient consent when discussing prescriptions.
- **Crowdsourcing to provide support for patients** – Patients register on a free online community to share experiences and symptoms. The online community allow a drug company to invite diabetics to participate in research into the efficacy of a drug to treat sight loss.

Some case studies perceived as higher value, or higher risk, to society than others.

We asked participants to **map the studies according to their perceived value to society or risk to society**.

---

[2] The case study 'Monitoring safety of drugs and medicine' was first introduced to participants as being conducted by a public health regulator. This was intended to contrast with the five other examples of commercial access so that the research could explore how responses might be different if a non-commercial organisation was involved. In follow-up discussions, the regulator was replaced with a pharmaceutical company and the question of commercial involvement was then further explored.

- Participants defined 'value' as having clear potential for patients, society and future **generations to benefit**. This meant *Linking data in the NHS* and *Monitoring the safety of drugs* were acceptable and seen as valuable. This public benefit could be created by a commercial organisation, but the organisation had to be trusted to deliver it.

- Participants **traded off value and risk by thinking about how the public stand to gain**. In examples where the public value or benefit was not clear, participants found it hard to weigh up public good against personal risk. The context collapse made these new data transactions hard to define.

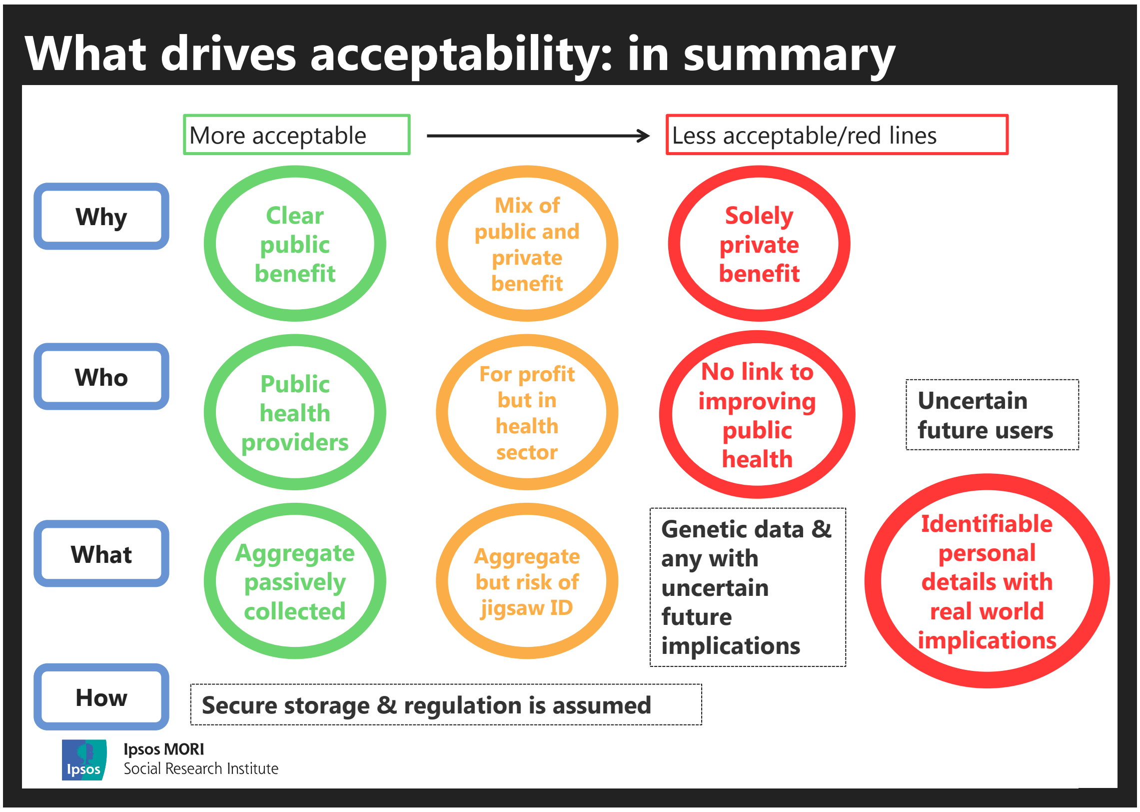
- **Genetic sequencing was considered to be the most risky example but also the most potentially valuable**; genetic data both most private, and most potentially valuable. Participants were concerned, overall, because thus the full extent of what might be possible with this type of data is as yet unknown.

Participants applied four 'key tests' to every example of data access

Participants discussed the issues arising from each case study as they saw it, and came to conclusions about how acceptable each might be. Their approach could be summarised as **applying four key tests** to all the case studies.

These tests are always applied in the same order. No matter how 'well' a project does on tests 2, 3, and 4, it will still be rejected if it fails test 1.

Figure 1.3 – The four 'key tests'; driving acceptability



What drives acceptability: in summary

More acceptable → Less acceptable/red lines

Why: Clear public benefit; Mix of public and private benefit; Solely private benefit

Who: Public health providers; For profit but in health sector; No link to improving public health; Uncertain future users

What: Aggregate passively collected; Aggregate but risk of jigsaw ID; Genetic data & any with uncertain future implications; Identifiable personal details with real world implications

How: Secure storage & regulation is assumed

1. WHY (Does the activity's outcome have a provable and sufficient public benefit?)

A **clear benefit both to individuals and to wider society** was seen as the only good rationale for breaking privacy, and this was the primary driver of acceptability for participants. When these benefits are perceived (and the organisation is trusted to deliver them) all participants in the workshops accepted commercial access to health data in principle.

The results from the survey showed that, **broadly, more support than oppose commercial access to health data for *research purposes*** (which might be seen as a positive social benefit). When this research is explicitly at threat, respondents are even more likely to back commercial access.

Red lines:

When no benefit to public health is perceived, commercial access is unacceptable to all in the workshops, except the most laissez-faire.

2. WHO (Can the organisations doing this be trusted to have public interest at heart?)

Healthcare professionals and general public participants are most open to the idea of **academic researchers, charities, or partnerships between these and the public sector** having access to health data. GPs also thought partnerships could be a safeguard against unethical conduct; both on the part of the commercial organisation, and to ensure the NHS is also working in the public interest.

Patients thought that **charities and universities** should do more with data, as they saw the public interest very strongly.

Beyond these most acceptable research, charity and partnership options, general public participants had **a hierarchy of acceptable commercial** organisations:

- Specialist analytics and research companies working closely with the NHS; participants did not know much about them but felt they were benign.
- Pharmaceutical companies. Some feared 'Big Pharma' had an agenda at odds with the public interest, but most acknowledged that a regulated pharma sector helped public healthcare and needed data to do its job. Because of the profit motive most people felt regulation was required.
- Retail and pharmacy sectors; motivated by profit but still producing a net benefit to healthcare.

Red lines:

- Participants in the workshops wanted **insurance companies not to have access to healthcare data at all.** Insurance companies were seen as detrimental to **individuals**. Participants said they always grudged giving insurance companies access to their personal records, as the industry is seen as having high charges but often refusing to pay out; the public believe this would get worse if more health data was available to the industry. Health insurance of all kinds was also seen as something that only affluent people would choose to have; the industry was perceived to be harmful to **society**, because it works against the principle of a public health service.

- **Marketing** companies were seen as **detrimental to individuals and a source of frustration**, plus motivated by their own profit, not public health, so producing private rather than public value from segmenting and marketing to consumers. People also believed that individuals would likely be subjected to direct marketing, narrowly targeted at them.

- Where the organisation's **purpose was uncertain** (for example a proposed future genetics data company) **participants mostly erred on the side of caution and would prefer data not to be shared.**

- This chimes with findings from the **quantitative survey, which found that low proportions of the public support insurance companies and marketing companies having access to health data.**

- **Third party access, passing data on to others beyond original use, was also a red line.** Participants did not want commercial organisations to profit several times from **re-selling** data, or even from analysis based on the data.

- Also, participants believed that no amount of security could ever totally remove the risks involved in sharing data such as leaks and hacking. **Allowing third party access was felt to increase these risks and create feelings of potential loss of control.**

As well as these red lines, participants asked for a particular regulatory safeguard: they wanted someone independent to prevent companies 'spinning' to enhance the social **worthiness of their activities in order to gain access to data; or taking it for one purpose but using it for another.**

3. WHAT (How anonymised and/or aggregated is the data?)

Participants accepted the description of aggregated data and did not find this data to be very risky to the individual. **Aggregated data could be risky to society, if re-purposed in ways that led to groups being discriminated against.**

Participants were concerned about the theoretical risk of jigsaw identification from anonymised individual data and assumed that if it was at all possible, it was a risk. HCPs and Cohort members understand anonymised individual data to have a low risk and are less concerned.

Some general public participants think we are **heading for a surveillance society and the collection and sharing of data should be minimised on principle.**

Red lines:

Anything that **risks personal harm**, especially to vulnerable individuals; participants found it hard to judge the likelihood of different types of data causing these harms.

4. HOW (Does the safeguarding, access and storage protocol reassure me that the data will be safe?)

Most participants believed that in today's world of health data-sharing, good data security and protection is a basic hygiene factor. Older and less tech-savvy people were least able to judge the security of digital storage.

In terms of safeguards, **there is little understanding of the status quo (many of the safeguards workshop participants asked for are actually in place already).**

The quantitative research also showed that there is no silver bullet in terms of safeguards, and that the **precise nature of the safeguard is not as important as the trust that comes with knowing there is a safeguard in place.**

The call for safeguards reflects a wider concern about governance of data and the need for a greater discussion of how to regulate the outcomes permitted to come from data sharing.

Participants asked for particular safeguards which, while not always practical, reflected underlying concerns for equity, transparency and independent scrutiny by bodies free from vested interest. These included capping profits and restricting third party access (to stop companies exploiting public resources for profit), and transparency in sharing results and publishing analysis (so that public benefit was made clear).

These findings resonate with the quantitative findings, which found that the most mentioned safeguards from a prompted list are:

- Data not being passed on to third parties;
- Names and personal information being removed from the data;
- Sanctions and fines if companies are found to misuse the data; and
- Storage of data in a secure facility.

The role of consent is complex and reflects the fact that many members of the public do not know how data sharing would work in practice:

- **Healthcare professionals** felt that if third party access was involved, then there would be more need for **opt-in consent.**
- Participants wanted **data sharing of genetic information to be opt-in only.**

- Deliberation days tended to start with the premise that consent should be explicitly sought at all times, but with more information, participants **shifted their view** as they identified projects that they wanted to go ahead, but would be impractical to consent.

- When asked explicitly about consent, **over half of respondents in the survey agreed they would want to be asked permission, even if this meant some research not taking place.** This shows that, for many, consent is still a key stumbling block to acceptance of commercial access.

Participants talked about how the process of consenting could be improved.

They would like:

- Education on aggregation and anonymisation, to build public trust.

- Regulators or future commercial data-sharers to be held to high standards in terms of clear, transparent online consent processes; not confusing tick boxes or small print which is never read.

- Healthcare Professionals such as GPs to be trained as gatekeepers to explain how research and consent works.

Overall, workshop participants felt that if they **knew more** about the processes and safeguards in place they might feel more **empowered**, and hence more **open and trusting** in the decision-making process around data collection and sharing (and may not, therefore, need to opt-in).

If the four tests are met, most people are comfortable with commercial access. However, some members of the public simply do not want this work to happen at all.

In the survey, a quarter (25 per cent) would rather research did not happen if commercial organisations had to have access to the data.

17 per cent of people say they would not want commercial organisations to have access to health data for research under any circumstances.

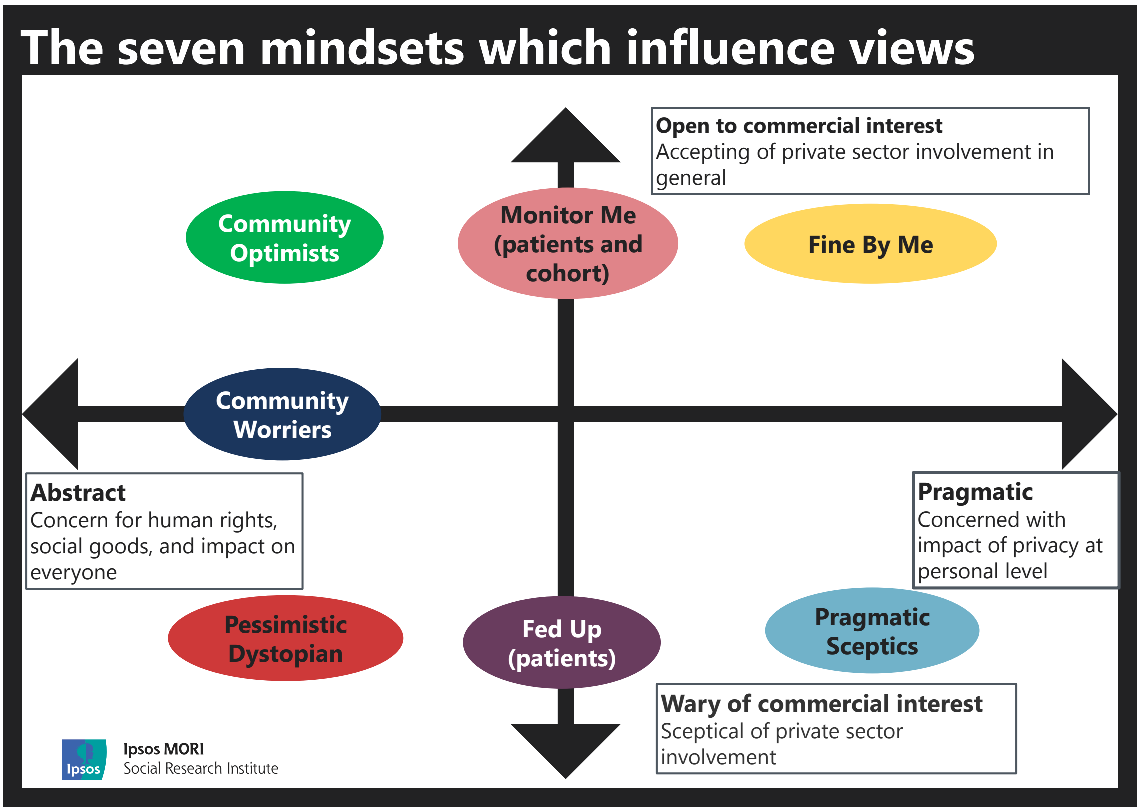
1.4 Mindsets: what makes a difference to views?

There are different reasons why individuals might hold an attitude. The non-linear relationship in the survey between age and acceptance of commercial access demonstrates that, while age is related to acceptance, there are other factors within the age groups that act as alternative drivers.

These sub-groups within sub-groups are not possible to tease out in a short questionnaire, but the **qualitative work allows us to identify differing attitudinal clusters that may cut across demographic variables.**

It seems that some concerns are underpinned by different beliefs about how society should operate and leads to different 'mindsets' about the idea of commercial access. We identified five different stances in terms of trust in commercial organisations in general, mapped against concern about personal and societal risks. Patients tend to fall into two further mindsets. These seven key mindsets are shown in figure 1.4.

Figure 1.4 – The seven 'mindsets'



1.5 Findings from the quantitative survey

- **More support than oppose health data being used by commercial organisations undertaking health research**. Over half (54 per cent) support commercial access to health data for research, while a quarter (26 per cent) oppose it. This leaves a fifth of respondents who either said that they neither support nor oppose commercial access to health data (19 per cent) or saying 'Don't know' (2 per cent).

- **Awareness of health data usage is low**. Just one third (33 per cent) have heard a great deal or a fair amount about how the NHS is using health data. This detailed awareness falls to 16 per cent for commercial organisations and 18 per cent for academic researchers.

- **Faced with losing out on research, people will opt for this research being done by commercial organisations.** A majority of people (61 per cent) would rather see commercial access to health data happen than lose out on the benefits that research involving these organisations can bring. A quarter (25 per cent) would still rather that research did not happen if commercial organisations had to have access to the data.

- Permission is still a key part of people's attitudes towards their health information. The desire for the NHS to ask permission before this kind of data sharing takes place is widespread, with 54 per cent saying that they would rather the NHS ask patients' permission before sharing data with commercial organisations, even if this meant that without permission the research cannot take place. The survey also found that a majority (53 per cent) want to see strict rules in place that data could not be passed to third parties.

- **Sharing health data for the purposes of insurance or marketing both face significant resistance from the public**. Just a quarter (26 per cent) support sharing anonymised health records with insurance companies so they can develop their insurance prices. Support for companies marketing health products using anonymised health records is higher, with 37 per cent supporting this purpose, but still lower than some of the support for commercial organisations generally using health data.

- **There is still a core segment who cannot see any circumstances under which they would allow commercial organisations access to NHS health data**. 17 per cent of people say they would not want commercial organisations to have access to health data for research under any circumstances. Of these, one-fifth (20 per cent) say commercial organisations cannot be trusted to store the data safely, and a similar proportion say that profit should not be made from health data, even if there are potential societal and health benefits as well.

- A "social" value to data resonates, but many still see a "financial" value to health **data.** The majority agree that health data has financial value (50 per cent) and societal value (67 per cent).

- **Having a safeguard in place makes a difference, whatever the safeguard is.** Using a split sample experiment, the survey found that there is no silver bullet safeguard that can restore public trust (between 56 per cent and 64 per cent agree with data sharing with any of the named safeguards in place). However, knowing that a safeguard is in place, regardless of the type of safeguard, makes a difference to acceptability of health data sharing with commercial organisations (agreement is 49 per cent for those respondents where no safeguard was mentioned).

- **Despite this, people are ready to insist on many different conditions to safeguard their data.** Of various different conditions that could be placed on commercial organisations to allay fears for the public, strict rules about not passing data on (53 per cent); all names and personal information being removed (52 per cent); and storage of the data in a secure facility (47 per cent) were the most popular. However, there is no clear preference for respondents on which conditions should be in place, and over two-fifths (43 per cent) want 3 or more of these conditions in place.

1.6 Communicating with the public

The deliberative work identifies key areas which will be core to any broader social discussion of the implications of commercial access to health data.

This will not necessarily create a public more amenable to data sharing; but a more informed and engaged public could contribute better to the debate on the role of data sharing in the changing context of data, healthcare and society

The areas are:

What different kinds of value does data have? Does data have long term value to society? If data is valuable financially, who should derive the value – the individual or state or business?

What should be the new social contract around health and digital data? With the rise of the ability to collect and use data, what should change about the way health services are paid for and delivered? What is the duty of the citizen, the government, and business?

Is our health data a 'common resource'? It is collected (sometimes) in people's most vulnerable moments. How should government act ethically, in relation to this valuable resource, in order to preserve public trust? The public do not want data to be used in the service of dismantling the NHS; how can this be prevented?

How can we situate conversations in the context of the future technology of healthcare? For example, how might wearable or other biometric technology shift focus towards individuals being in charge of their own preventative medicine? What do the public think about the 'quantified self'? How should social media or passively collected data be best used in healthcare?

Some of these issues are being discussed already in a range of studies reporting in 2016; there will be a need to draw shared learning together.

1.7 Conclusions

Table 1.1 – Conclusions and recommendations

Overall conclusions from qualitative and quantitative research	Recommendations for next steps
Most of the general public tend to accept commercial sharing of health data, as long as the four key tests are met.	If commercial sharing is on the table, the four key tests need to be applied upfront, so that the public can be reassured if they are asked to support this new way of using data.
Safeguards help the public feel reassured; the most convincing safeguards are those which regulate the profit motive in the interests of public benefit and create independent scrutiny and control. While safeguards on data handling are important, no one safeguard is a 'silver bullet'; the survey reveals that any safeguard is reassuring but no particular sanction or storage safeguard is more reassuring than any other.	Safeguards should be put in place first and be designed to enhance public control, for example opt-outs wherever possible, and overall designing a system where there are no hidden incentives for companies to behave badly.
There is a core group of those who do not want health data to be shared at all (17% do not want data shared for research under any circumstances; 25% would sometimes rather research did not go ahead than data is shared).	Policy and research interests are not likely to be able to sway this group. Opt-outs should be offered, along with clear communication about the safeguards and purposes of sharing, to minimise concerns among this group.
There are different views about different types of organisation, with some considered more acceptable than others for accessing data. • **Insurance is considered unacceptable** in the qualitative research and only 25% support it in the quantitative survey. • **Marketing was considered broadly unacceptable in the qualitative research, except in healthcare contexts,** and 38% supported this in the quant (NB a healthcare context was given). • **Third party access** to data was considered to be risky and not socially beneficial. The public do not want profit to be made from this resource without a company having a very explicit public benefit inherent in its work.	There is a need to identify and communicate a clear public benefit associated with data access if a company is involved, so that the public are reassured that profit motives will not override public benefits. To reassure the public, insurance and marketing uses of health data should not be allowed. There is a need to restrict third party access and companies redeploying health data for further profit, and to open a broader debate about what value data has, and to whom should accrue that value.

Overall conclusions from qualitative and quantitative research	Recommendations for next steps
The public know little about some key areas: • Not aware of the current range of **uses of health data**, beyond in their own care • Underestimate the **amount of data** currently collected and used in healthcare • Do not understand why the NHS would need/want to allow commercial access to data, do not know **how the commercial sector contributes to healthcare currently** • Little understanding of the **status quo when it comes to safeguards** (some that participants want are already in place) • Confusion about the **specifics of data and data science in general.** (e.g. the difference between anonymised versus identifiable data, or definition of aggregate data).	These point towards a need to engage the public further, inform and communicate, in order to have a more informed social debate on the uses of health data. For example, there may be a need to inform the public as to how statistics work, what data is, and how it is combined into datasets, again in order that they can be informed on the subject. This will put the public in a better position to understand the real risks or benefits of data sharing.
There are many different mindsets and perspectives on commercial access to data; views of data sharing are influenced by opinions about society and commerce generally. New technology has given rise to new ways of collecting data, both actively and passively; and new ways of using the data to create knowledge. This has led to blurred lines between traditionally private and public sector ways of collecting data, causing a Context Collapse. The public are finding it difficult to navigate 'data contexts' as they shift and change, and are therefore very wary.	Codes of conduct may need to give explicit and separate consideration to the needs and fears of different groups of the public, for example ethical frameworks could be constructed which set out the spectrum of acceptability for different publics, as well as ensuring that all bases are covered for everyone. The public will need help to negotiate the context collapse, 'read' different data contexts more skilfully and feel more able to make decisions. NB: this will not necessarily lead to support for commercial access to data, but more information may mean members of the public feel better able to protect themselves.

Overall conclusions from qualitative and quantitative research	Recommendations for next steps
Overall, this project provides an evidence base to influence government and to use public views to improve the process of data sharing.	Policymakers and the research community should take into account the findings of this report when designing new processes and policies.
There is scope for a wider ongoing discussion. A well-designed and timely discussion between policymakers, experts, research and clinical communities, and commercial organisations, as well as involving the public, could well shape the future of biomedical research, healthcare and notions of citizenship.	

2 Introduction

2 Introduction

This chapter sets out the project objectives and design and gives guidance on how to read this report.

2.1 Background

Data is collected throughout the health service in increasingly large quantities, as well as in the contexts of biomedical and health research. Data is collected for direct care and for secondary uses as well.

As sophistication in using large datasets increases, it is clear that commercial organisations can use our health, biomedical and genetic data in a number of different ways. There is interest in health data from a wide variety of commercial bodies, and a consensus that there may be various kinds of **value** in reusing data and allowing commercial bodies access (for example, economic value, or social value in terms of improving access to or design of services).

However, recent developments in health and more widely show that public trust in organisational use and handling of data is at a very low ebb. The public and healthcare professionals raised their concerns over the care.data programme, highlighting significant barriers to health data being accessed and reused – particularly by the commercial sector.

Over recent years, quantitative and qualitative studies, and public dialogue, have surfaced some of the public's concerns around commercial access to data. Ipsos MORI's ONS/ESRC-funded qualitative study, *Dialogue on Data*[3], found the public felt a **lack of control** over their own data and a feeling that data reuse was an **invasion of privacy**. While the study looked at researcher access to data, views applied to both the public and private sector. Further Ipsos MORI research for the Joseph Rowntree Reform Trust showed that 93% of the public think it is essential/very important to maintain the privacy of medical records[4].

In particular, Ipsos MORI's work for the Royal Statistical Society[5] (RSS) suggested that health records being shared with private companies is a scenario that, in principle, most people think should *not* happen. In awareness of this study, the Wellcome Trust wanted to investigate the nuance behind these findings.

There are many different types of commercial access going on. The Wellcome Trust therefore wanted to investigate how, and in what ways, the public would distinguish between them, and

---

[3] Ipsos MORI/ESRC/ONS (2014), 'Dialogue on Data', https://www.ipsos-mori.com/researchpublications/publications/1652/Dialogue-on-Data.aspx

[4] Ipsos MORI/JRRT (2014), 'Privacy and data sharing', https://www.ipsos-mori.com/researchpublications/researcharchive/3388/Poll-on-privacy-and-data-sharing-for-The-Joseph-Rowntree-Reform-Trust.aspx

[5] RSS/Ipsos MORI (2014), 'Public attitudes to data sharing', https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-lessons-for-policymakers.aspx

whether the type of data used, and the types of data user, would have an influence on the level of acceptability of commercial access.

In June 2015, therefore, The Wellcome Trust commissioned Ipsos MORI to conduct a rigorous study investigating attitudes towards commercial access to health, biomedical and genetic data.

The study involved both deliberative workshops and a follow-up quantitative survey.

The qualitative study was designed to investigate, in detail, views of a range of real-life examples, drawing out the implications of different variables on acceptability. Hypothetical examples were also explored, in a structured fashion.

The emerging findings were then used as a start point to develop questions for quantification. Ipsos MORI conducted a face-to-face survey of adults in Great Britain to collect quantitative findings about their attitudes towards commercial access to health data. The survey contained a range of different question approaches to tease out the impact of different variables upon views.

2.2 Objectives

The overall objective of this research was to understand how attitudes towards commercial access to health data are formed and influenced, among a reflective cross-section of the British general public and specific audiences such as healthcare professionals, patients, and members of cohort studies.

The primary objectives of this research were to:

- Identify **factors influencing attitudes** towards commercial organisations accessing health, biomedical and genetic data;
- Identify **governance, safeguarding and communications actions** that could help improve trustworthiness of research uses and protections of data; and enable public trust in access to data to be developed over time;
- Provide an **evidence base for the Wellcome Trust** to draw upon as it considers how best to develop policy and engage other relevant bodies, research teams and the public on these important issues.

Secondary objectives were:

- In order to best understand the detail of the factors influencing attitudes, to explore how people perceive the relative importance of different **variables** in sharing health data with commercial bodies.
- Identify any emerging **differences** between sub groups and different audiences.

- **Quantify** views on key topics which emerge from the qualitative phase.

2.3 Project design

The Wellcome Trust identified four key audiences to be consulted separately for the research; the general public, patients, healthcare professionals, and cohort study participants. During September and October 2015, Ipsos MORI conducted **sixteen qualitative workshops** across Great Britain. These included:

- Eight full-day deliberative workshops with members of the **general public**;
- Three evening workshops with **GPs and hospital doctors**;
- Four evening/daytime workshops with people who have a **long-term condition**, including rare diseases;
- One full-day workshop with 'Children of the 90s' cohort members.

Further details on the workshop audiences and fieldwork locations can be found in the Appendix.

The **general public** reflected the largest population and understanding their views was key since they were less likely than other groups to have considered the issues (e.g. personal/public impact), despite being affected. **Healthcare professionals**, **patients and cohort study members** were identified as important *'special public' audiences* to consult, due to their increased contact and familiarity with both the health service and health data-sharing activity. It was considered that their personal experiences might have afforded them a more nuanced or complex perspective on the potential benefits and risks of health data-sharing and commercial access.

Participants in the general public workshops were recruited on-street by specialist Ipsos MORI qualitative recruiters. A mix of urban, semi-urban and rural locations was chosen to ensure good geographical representation, and recruitment quotas were set to ensure that overall people of a range of ages and from a variety of ethnic and socio-economic backgrounds took part.

Healthcare professionals were recruited via Ipsos MORI's online panel and included participants with a range of number of years' experience and qualifications.

Healthcare professionals were included as an interesting audience to speak with given their relative closeness and insight into existing health data-sharing practice. They are also controllers and collectors of patient data so would be gatekeepers to data sharing in general. It was thought that they were likely to have a better grasp of the potential benefits of commercial access to health data, but that they might also have concerns given their professional capacity, such as the impact on doctor-patient relationships and the administrative demands that increasing amounts of data-sharing create.

Patients with long-term conditions were mostly recruited through a standard on-street recruitment method by asking people whether they would describe themselves as having a long-term illness or disability and the extent to which it affected or limited their day-to-day activities. **Patients with rare conditions,** and their carers, were recruited via the charity Rare Disease UK[6].

Patients with long-term conditions and therefore greater reliance and interaction with the health service, could potentially have had a different view on the issues. Whether this hypothesis was borne out or not, their perspective was important to capture, as they are an audience particularly impacted by sharing and accessing health data. Patients are potentially the greatest beneficiaries of sharing health data, but also the most vulnerable and easy to identify, especially those with rare conditions. Thus it was important to understand the beliefs and fears patients might have and whether they are more or less in favour of commercial access than other groups. Another reason to split out this group from the rest of the general public was in case some patients were living through the implications of health and data-sharing, making it a sensitive topic for them to discuss.

Finally, **cohort study members** were recruited by the 'Children of the 90s' research study team at Bristol University.[7]

Cohort study members were identified as a key group to consult due to their greater experience of sharing medical and health data - in particular of doing so for the purposes of scientific research carried out independently of the healthcare system. It was felt that this group might have quite a unique perspective, beginning from a more informed and knowledgeable standpoint and having had greater exposure to the potential consequences, positive and negative, of data access. The Wellcome Trust was particularly interested in conducting research among this group so that they could identify effective ways of supporting their research communities in future when those communities engage with research participants on issues surrounding data access.

2.3.1 Deliberative workshop approach

A deliberative workshop approach was taken due to the complex nature of the issues around data use and reuse, and the low levels of awareness and understanding with which many people approach these issues. A workshop is an ideal, **open environment** that gives people time and space to learn new information, ask questions, change their minds and develop their views with others like them. Workshops also allowed sufficient time to explore a larger number of variables via case studies and other stimuli so that participants were able to see how data-sharing currently operates in the healthcare and research system.

2.3.2 Design of materials for exploration: case studies and question design

The Wellcome Trust convened an Advisory Group to provide a sounding board and critique for the design of the study. Other stakeholders outside this group also contributed to the

---

[6] http://www.raredisease.org.uk/
[7] www.bristol.ac.uk/alspac

development of materials. We have listed membership of the Advisory Group and other contributing stakeholders in the Appendix.

The Advisory Group met before and during the study to help scope design and to hear and comment on initial presentation of findings to refine this report. The group also contributed to the development of materials by reviewing the case studies which formed a key part of the qualitative study.

The Wellcome Trust and Ipsos MORI collaborated to develop a set of detailed case studies. **Six case studies** were designed to explain to participants, in simple English, why companies access data, what their aims are, how they go about it and what type of data is accessed. In general public workshops, all six case studies were discussed; in shorter sessions with patients and healthcare professionals, four were selected.

The studies formed a set of examples to inform research participants about the scope and variety of commercial (and non-commercial) access to health data that currently takes place. A case study approach was chosen in order to go beyond what past research on commercial access to health data has shown, by encouraging people to engage with what happens *in practice* and use factual accounts as a starting point for discussion rather than media reports or hearsay.

Discussion of the case studies was preceded by a general, **high-level discussion** of what data means to people, and what they consider health, medical and genetic data to be, as well as their awareness of and feelings towards current practices of data use and sharing. We also explored where people found their ideas about data sharing, as the Wellcome Trust wanted to identify whether and how opinions are formed as a result of media reports and other exposure.

Examples of real-life data sharing in the healthcare and research system were selected for their ability to **inspire discussion** of a range of different variables which might affect public acceptability, such as the different types of organisation involved, different types and formats in which the data is stored and analysed, and different the types of safeguards in place to protect against incompetence or misuse.

Each current example was then followed-up with a series of hypothetical 'what if' scenarios, designed to amplify and extrapolate certain (often controversial) elements of the data-sharing activity, and push participants to consider what might happen in future. Participants were forced to weigh up public value against private benefit, altruism against self-interest; and in doing this they discovered where they draw 'red lines' in data access and what governance and safeguards, in their opinion, any commercial access in future would require.

The **order of presentation** was rotated so that the different examples could be assessed on their own merits rather than participants being conditioned by their cumulative responses to previous examples. Across the workshops, each case study had the chance of being assessed first and freshly.

After discussing each of the case studies in turn, participants were asked to assess, first by themselves and then as a group, the relative value and risk of each against the others and to plot them on a grid with 'Value to society' pitched against 'Risk to society'.

This exercise was followed by detailed discussion of the **rules, regulations, safeguards and sanctions** that might, or might not, make a difference to views of commercial organisations' access to health data.

See appendix for the case study and 'what if' slides and the full discussion guide for the general public groups (other guides were variations on this and are available on request).

2.3.3 Improvisation team and digital anthropologist involvement

In two of the general public workshops (those conducted in Wrexham and Sutton Coldfield), an improvisational acting team performed a series of sketches at various points throughout the day, to help bring to life the case studies and highlight emerging points of view, as well as providing stimulation for plenary discussions. Lydia Nicholas[8], an anthropologist who specialises in digital constructions of identity, was also present at the Sutton Coldfield general public workshop, observing the discussions in order to provide further insights about the language people use and how they conceptualise themselves in relation to the data they share. Lydia wrote a briefing paper for Ipsos MORI and discussed her findings, which helped to illuminate and enrich analysis and the framing of our reporting. We have indicated specific references where relevant, in general Lydia's findings are woven throughout.

Figure 2.1 – Participants watching improvisers

---

[8] http://www.nesta.org.uk/users/lydia-nicholas?gclid=CL_vxf2mjcgCFRATGwodGXYOkg

In terms of **methods**, in this qualitative study the use of **improvisers** was very useful and can be recommended:

- To enhance the emotional weight of discussion and foreground important points. Other applied theatre techniques such as Playback Theatre could also be recommended for future ethical discussions, to situate real experiences in the context of principles and allow facilitators to open up more abstract debates. This could also be helpful to discuss issues with patients or those with a more emotional connection to the subject.
- To point out to people where ethical issues are at play by creating laughter, surprise and unexpected juxtapositions which could be used in the discussion. Other ways to do this could have been to add observers who also feed back to the group, for example ethicists or practical philosophers, who can draw out the broader social questions in play at different times within the discussion.

2.3.4 Quantitative study

The qualitative work was followed up by a quantitative survey that examined some of the issues arising from the workshops. The questionnaire was designed in conjunction with the Wellcome Trust and the advisory group, and aimed to fill in gaps that previous surveys have left open on the subject of data sharing with commercial organisations.

The design of the questionnaire took into account interesting parts of the qualitative findings that appeared to warrant and suit quantitative follow-up. However, at the same time, it was appreciated that priorities needed to be made about what could be asked in the allotted time, and that the exercise of conducting quantitative research is very different from a qualitative project, where participants are introduced in depth to the context across an entire day.

2.4 Reading this report

When reading the qualitative chapters, the following note may be helpful. Qualitative research approaches (including deliberative workshops) are used to shed light on ***why*** people hold particular views, rather than how many people hold those views. It is used to explore the nuances and diversity of views, the factors which shape or underlie them and the ideas and situations in which views can change. The results are intended to be illustrative rather than statistically reliable. Given the qualitative nature of the data collected from the workshops, this report aims to provide detailed and exploratory findings that give insight into the perceptions, thoughts and feelings of people, rather than statistical evidence from a representative sample.

It is not always possible in qualitative research to provide a precise or useful indication of the prevalence of a certain view, due to the relatively small number of participants generally involved (as compared with the larger respondent bases involved with quantitative studies). So, the views of proportions of the qualitative group should not be extrapolated to the population at large. Sometimes, ideas can be mentioned a number of times in a discussion, and yet hide the true drivers of thoughts or behaviours; or a minority view can, in analysis, turn out to express an important emergent view or trend. The value of qualitative work is to identify the issues which bear **future investigation**.

Therefore we use different analysis techniques to identify how important an idea is. The qualitative report states the **strength of feeling** about a particular point rather than the number of people who have expressed that thought. Having said this, is it sometimes useful to note which ideas were discussed most by participants, so we also favour phrases such as "a few" or "some" to reflect views which were mentioned infrequently and “many” or “most” when views are more frequently expressed. Where views apply only to a subset of participants, e.g. participants in Swansea, we have highlighted this in the text, as this may indicate differences by rurality, for example. Any proportions used in our qualitative reporting (e.g. a ‘couple of’ or ‘a few’ participants), should always be considered indicative, rather than exact.

Verbatim comments have been included in this report to illustrate and highlight key points, i.e. those views either shared by a large number of participants or reflecting the strong views of a smaller subset. Where verbatim quotes are used, they have been anonymised and attributed by location and group/workshop type (e.g. General Public, Dundee).

The **quantitative** chapter presents the headline findings in chart or table format and then describes these findings in the text, going into any details of sub-groups. Only selected statistically significant differences in sub-group responses are reported upon. Please see the appendix for details on statistical significance. The quantitative findings are also used at various points in the qualitative chapters to add extra insight throughout.

The survey results can provide an indication of the **prevalence of views in society**, and also about how different groups hold differing attitudes. This can be useful in understanding which factors might affect a certain point of view. These should also be viewed with caution – sub-groups do not tell the whole story. Age differences, for instance, might be a function of an entirely different set of factors, like location or internet use.

The qualitative findings do not exactly map and mirror the quantitative ones. Workshop participants went on a much more substantive journey through the day and their views were nuanced. In the qualitative work, there was more scope for getting information about details while in the quantitative survey there was no discussion or context given. We should be cautious, therefore, in drawing too many conclusions or forcing both pieces of research to tell the same story.

The report links the two, but they stand alone. We have indicated where parallels are striking.

Data: singular or plural now?

In order to follow the progress of language change, and to write in a simple way, we have chosen to say ‘data is’ rather than ‘data are’ in this report.

2.5 Acknowledgements

The project teams at Ipsos MORI and the Wellcome Trust would like to thank all who provided help and support in the design and delivery of this research, in particular: digital anthropologist Lydia Nicholas; improvisational actors James ‘Lloydie’ Lloyd and Liz Peters; all members of the Wellcome Trust’s External Advisory group, and individual stakeholders who contributed to the design of the case studies and other research materials; staff at Rare Disease UK for their assistance with recruitment of patients and carers of patients with rare diseases; staff at Bristol University/The Avon Longitudinal Study of Parents and Children (ALSPAC) for managing recruitment for the workshop with cohort members, and assisting in the smooth running of the day. Refer to the appendix for a detailed list of contributors.

3 Low awareness of data and new contexts around data sharing

3 Low awareness of data and new contexts around data sharing

This chapter provides the cultural and social context of the discussion of commercial access to health data.

Participants have limited knowledge initially about data, about anonymization and aggregation of data, about the regulations surrounding it, and about the role of private companies in the healthcare system. In the deliberative work, this led to wariness and scepticism about the idea of commercial access to healthcare data.

Importantly, contexts around information and data are generally changing, owing in part to new technology for collecting and sharing data. Expectations of privacy and ownership are changing, as new types of data collection usher in new norms. These norms, however, are not universally agreed yet.

The situation in flux also creates a sense of confusion and wariness around data sharing. Overall, the fear could be summed up in the idea that information sharing seems unidirectional, shadowy and hard to understand.

"It's a one-way mirror; they know everything about you, but we don't know what they're doing with that."
Patient (severe conditions), London

3.1 What is data?

Workshop discussions began with an exploration of what people think 'data' is and what they understand by the terms 'health data', 'medical data' and 'genetic data'. Initial top-of-mind association exercises were followed by a high-level discussion of different ways in which data can be actively or passively given, collected, used and shared.

Participants first gave examples of the information they shared about themselves *as individuals* – name, address, date of birth, National Insurance Number, medical history, insurance details, credit history. They tended to see this type of data as 'my data', usually because it is **sensitive and identifiable.** Going even further, for some people this was not only 'my data' or 'data about me' but simply 'me'. For some, 'I am my data'.

"Name, address, phone number, email, diagnosis, history, age, sex."
General public, Swansea

"I wrote identity which is really broad – ***once you've given them all those things they've got 'you' – they've got an image of your identity. There's nothing you can*** *do to prevent what they do with it."*
General public, London

Participants did not make a clear distinction between 'health' and 'medical' data. Across the workshops the same examples were given for both: illnesses; medications; visits to the GP or hospital; blood pressure; BMI; smoking and drinking habits. Participants conceived of health and medical data as individual level data and had similar concerns about using and sharing it as they did with other types of data.

Most participants (aside from those with rare diseases, some of whom knew more) had only **a vague understanding of genetic data** and treated the subject with ambivalence. It was often associated with fingerprints, family history and inherited traits or illnesses and for some raised more sinister associations with cloning. Genetic data was seen as inherently personal and much **more private** than other types of data. Participants knew it was a new type of data, which potentially could be used to shed light on important areas of life which have so far been mysteries; why one person falls ill when another is well, why one person has children and another is infertile, why some have allergies, some become obese, and so on. The idea of this potentially open-ended set of information, so personal to 'me', being known, often led to anxiety towards genetic data being held, used and shared at all.

"It's creepy. They know more about you than you know!"
General public, London

Some, notably younger participants, were awed by the idea of genome "scanning", seen advertised on television.

"That's really cool – to have your whole genome scanned! If it's pulmonary or *cardiovascular then you can do something about that."*
General public, London

Participants' top-of-mind worries around data tended to focus on the individual consequences and harm which could result from data getting into *"the wrong hands"*. Many raised concerns around **the detrimental effect that certain information** – **medical conditions or health behaviour** – could have on, say, an individual's employment prospects or financial status. This was a particular worry among patients with long term health conditions.

"They do penalise people with illnesses. [Like] car insurance. My ***daughter's a*** *prime example –* ***she's got MS but she's got to have her driving license renewed every three years and she pays a higher premium. She's got MS but she's as*** *capable of driving as anyone."*
General public, Swansea

“The biggest problem with that is insurance companies – they take your data, use that against your children.”
Patients (rare conditions), Sheffield

“Sometimes when people know you have an illness they will use it against you.”
General public, Glasgow

The question of **who owns data** was one that people returned to during the deliberative discussion. Participants were not consistent in how they referred to data, sometimes referring to ‘my data’ and other times being less specific; this highlights their uncertainty in pinning down who owns data, and furthermore who should profit from it. In general, though, they believed they owned their personal, individual-level data, which related to health and in other arenas too. They assumed data was usually kept private.

"But the question is who ***owns that information… because if it’s about me, it’s my*** *information."*
Patients (rare conditions), Sheffield

There is some indication that the norm of automatic privacy might shift in future, in certain contexts; see section 2.4 below. Nevertheless, in this research **participants took as axiomatic that data about people and their health should and would be kept private unless there is a good reason for sharing it.**

Healthcare professionals had a more nuanced view, pointing out that the patient could own the data, yet still not have the right to opt out of having the NHS use the data. The onus they see is on the NHS to keep the data safe.

"Any data which is provided by or comes from the patient belongs to the patient, but at the same time if they have given that data, then permission goes to the NHS."
Hospital doctors, Birmingham

3.2 Low understanding of aggregation and anonymisation (but better **understanding of 'statistics')**

What data were we looking at? The materials and examples used in the discussions related to data at different levels of anonymization:

- Identifiable data
- De-identified or anonymised individual level data
- Aggregate datasets
- Information about aggregate data for example averages and percentages based on the data

When asked to think about health and medical data that can be used in **aggregate**, the public spontaneously mentioned statistical data, such as A&E attendance figures or the number of people living in a certain area with heart disease, and how it could be used to improve services. The term statistics was broadly understood, but aggregate was not.

***"Anything that's held to help** – for example to target diabetes services in an area where there are lots of people with diabetes, and things like that – **that's got to** be a good thing."*
General public, Sutton Coldfield

Many participants had a limited grasp of the value of aggregate data and the opportunities it offers for understanding trends and patterns in human behaviour, health needs and treatment. They had little practical knowledge of the processes involved in transforming personal, individual-level data into an aggregate dataset that could be used to generate these insights.

"It says so they can predict what will make you ill or better. How? Are they god? How can they work all that out?"
General public, Glasgow

***"They don't tell us what will happen to the information, what they're going to do with it or where it's going – you just fill in a form and then you don't hear about it** again."*
General public, Sutton Coldfield

A poor grasp of statistics, the principles of aggregation, and no awareness of **data science** meant most participants did not see data – either at individual or aggregate level – as having financial value or other social value. Hence they did not think in terms of what they might be offered in return for their health data.

"There are benefits if it can be used in a useful way – but you get personally ***targeted. You can't see the changes to the waiting times. You don't see how it's*** *got anything t****o do with you, but you just see when you're being targeted."***
General public, London

Participants did not use or really understand the terms 'individual level' and 'identifiable'. For some, once data was anonymised it no longer felt like 'their data'. Others lacked understanding of, or trust in, anonymization, and also did not know how data is actually held in the health service; therefore they still felt that the data would be connected to them and had concerns about the potential impact on them.

"You can request your medical records now – ***the fact you can do that means it's definitely being held somewhere, and that it's not encrypted. If I go to a hospital 40 miles away they can get my records, so it's on a system somewhere."***
General public, Belfast

"I don't really care about my information going out to improve services, what I care about is people knowing whether it's me or not. If the marketing or insurance *people know that in a particular postcode area that a number of people have one leg and three* ***cars, I'm fine with that, as long as they're not throwing spam and I'm not getting junk mail. If I want a service I want to be able to find that service."***
General public, Swansea

Some, particularly patients with rare diseases, worried that even if anonymised they could still be identified due to the small numbers of people with their condition.

Even among those who did understand the opportunities they mainly thought that the **outcome of data sharing would be marketing and sales targeted at the individual**. Familiarity with this kind of data-sharing can colour views of health data sharing and mean people may initially struggle with the idea of more socially beneficial ends.

"Sometimes I think 'how do you get my number' - you get people ringing you about PPI, or saying ***you've had a car accident."***
General public, Wrexham

Overall, participants tended not to see a link between aggregate data and the information they each share when they interact with the healthcare system, or how it can be joined up.

This is a knowledge-gap that was more common among older less technical and digitally-savvy participants. These groups tended to raise simple questions about data collection, for example how could they ever be certain that all personal details from a data file had been removed. They were often unfamiliar with manipulating data using Excel and they latched onto the idea that the data file contained information about *them* and it therefore felt like *their* data.

This starting point explains why so many reassurances were required when data was proposed to be shared.

On the other hand, participants **had few concerns about allowing commercial companies access to aggregate health data which is *passively* collected,** such as hospital waiting times figures or number of hospital admissions. They associated this type of data with statistics and percentages and did not discuss it in terms of 'my' or 'our' data since it bears no relation to them personally, and carries little perceived risk. Some felt that sharing this data with commercial companies could still pose a risk to society – see the discussion of acceptable purposes for data sharing in the next chapter.

In the workshops, it was often mentioned that *"people are not numbers"*.

> *"I might be part of that group who has a heart attack, but I might be different – **you don't really know me, you don't know who I am**."*
> General Public, London

Even when statistical evidence makes it more likely that an individual might form part of a certain group, there is great resistance on an emotional level to individuals being 'pigeon-holed' in this manner or their access to services being determined by this. To some extent this shows a lack of knowledge of how mass services are run, based on large datasets, but also illustrates a real feeling that social decisions should be made based on something more than the aggregation of people into datasets.

3.3 Low awareness of how healthcare and biomedicine work today

Participants did not have a complete understanding of how the health sector works and how services are delivered. From early on in discussions, participants displayed **little understanding of the role that commercial companies currently play within healthcare and biomedical research**. This was the case across general public workshops, age groups and types of people.

General public participants overall tended to think about the NHS providing most services and its being funded through general taxation. They did not spontaneously reference the role of academics or charities in biomedicine or healthcare. They also assumed that commercial companies were not involved in delivering health services, nor in research which contributed to the sector, nor in the analysis of health data.

> *"Private companies have no need to have my medical information."*
> General public, Glasgow

Participants also wanted it known that mostly, in principle, they would **prefer the NHS to retain all its functions in-house rather than allowing private sector involvement**. This was the case even for those who knew there were private sector companies currently delivering services.

"The NHS can't afford to do (this) ...private healthcare is the reason the NHS can't. Private healthcare is flourishing. It's wrong."
General public, Dundee

Older participants tended to hold these views most strongly, alongside worries that the NHS was being privatised.

"Don't you think the money they're charging could go on the NHS rather than having [an independent] body [doing analysis]?"
General public, Sutton Coldfield

"It's like they're saying to us that they're gonna help us but they're gonna make money out of it as well. I'm not sure..."
General public, Dundee

Participants spontaneously drew no distinction between the private sector providing data analytic services to the NHS and the NHS in general 'being privatised'.

"The National Health Service is the National Health Service. Not the Private Health Service. They don't need to know what the data is."
General public, Swansea

Furthermore, there was low awareness of existing research processes; people did not know about the roles that **universities, charities, sponsors** and commercial organisations can have in health and medical research.

"If they're doing research you imagine them in a university lab don't you."
General public, London

Despite low awareness among general public participants of how medical and scientific research is actually carried out, participants tended to **respond positively to the principle of "pure research";** albeit they thought of all research as downstream and applied research to find cures for diseases[9].

This lack of awareness of the current state of affairs led to an initial sense of surprise and often a negative knee-jerk reaction to the very principle of commercial access to data. **The participants for whom the whole idea of commercial involvement in healthcare was** shocking 'new news' tended to express stronger concerns about commercial organisations having access to health data, and to stand by these views as they deliberated.

"(Is this going to) the whole industry? Drug companies, areas I don't know about?"
General public, Dundee

---

[9] We know from other public dialogues on basic bioscience that the public find blue sky research less easy to understand and value than they do applied research; see e.g. our recent report at http://www.babraham.ac.uk/get-involved/partnerships-page/public-dialogue/final-report

In some workshops, participants said that they almost felt **deceived**, that they now heard lots about private sector involvement and data-sharing taking place without this being generally known.

"You feel a bit in the dark."
General public, Dundee

This sense of unease is compounded as participants had very low awareness of **how data is currently used within the healthcare system**. Their initial associations in the workshops with data collection and data sharing focused on individual care, and individual-level data. Participants said they did not want private sector data sharing, however they are assuming this to be individual-level data:

"I think that's against the law. Everything's supposed to be confidential."
General public, Sutton Coldfield

""It's never really made clear how much of that information is being stored with my name against it."
General public, Belfast

Conversely, individuals wanted data *within* the NHS to be shared more widely for better care. Participants mentioned visiting their GP or hospital, and medical records being used and updated during consultations. Some described personal experiences where data should have been shared, but was not. This led them to question the efficiency of the whole system.

"I'm always having to explain to each (of the specialists I see) what's happening with the other. I don't think they're sharing anything."
Patients (severe conditions), London

"I'm a bit concerned about hospital records. I had a knee replacement that didn't go to plan and it fathoms me [sic] the things they know about you, you tell one hospital something and you go to another and they know all about you..."
General public, Sutton Coldfield

"The surgery called about a patient with the same name as me and that scared me – they could pull up the wrong medical record... They tried to give me a prescription that was not me – ***it scared me because you've given me a prescription for somebody that's not me."***
General public, London

3.4 Low understanding of safeguarding practices, and the reasons why data would need to be shared anyway

There was no awareness of the current regulatory framework around the use of anonymised health data. Few had heard of the Information Commissioner, and there was a sketchy knowledge of the role and remit of the Data Protection Act.

Participants showed little knowledge of safeguarding practices, as this assumes a background knowledge of how (or *if*) anonymisation happens, certain statistical concepts or of the different types of dataset that are publicly available. This meant that general public participants assumed that they **would be able to opt out** of any data sharing; there was no understanding that anonymised health data usage is currently allowed.

“I feel opt in should always be the option. Nobody should assume that you should need to opt out. Always opt in. That choice is taken away from you. Not everyone is going to be that clued up.”
General public, Glasgow

Furthermore, while people can imagine why the *NHS* itself might want to collect and analyse aggregate health data, they initially struggled to see why an external, commercial company would want to look at it.

Healthcare professionals and patients with lots of contact with the healthcare system and/or multiple experiences of sharing data, do appreciate why the data might be required. However they raise the issue of the quality of the data. These groups are sensitive to the inherent subjectivity involved in health data collection and the responsibility that GPs hold to do this accurately.

*"I want my doctor to review my condition, because it alters really fa*_**st. I’m already**_ *dealing with 3 hospitals and none of them talk to each other. My GP is the only one of all of them likely to have all my information. She needs to be reviewing* ***what happens with my drugs. Usually my GP has to go and look for it as I’ve told her! Maybe if I’ve done something and haven’t updated things with my GP for 2 weeks, it would fall between the stalls, the pharmacist wouldn’t know."***
Patients (severe conditions), London

“There’s a variance between clinicians using coding *– if clinicians report then it* ***will come up, but if they don’t it won’t*** *[…] that affects the analysis. Garbage in, garbage out.”*
GPs, London

The survey demonstrated that detailed awareness of how the NHS uses health data is low, with just a third (33 per cent) reporting to have heard a great deal or a fair amount about how the NHS is using health data. This awareness falls further still to just 16 per cent and 18 per cent awareness of health data use by commercial organisations and academics, respectively.

There is a **link between low awareness and understanding, and concern about commercial access** to health data. The survey results found that those respondents with a low awareness of health data usage and, related to this, a low educational attainment, tend to be less likely to support commercial access to data.

3.5 Changing expectations of privacy: a 'Context Collapse'

Participants identified four broad situations in which they give or share data. Examples covered both actively and passively given data, as well as one-off and longitudinal forms.

- Purchasing and commerce – for example purchasing patterns recorded through a store card, online purchasing history;
- Social media and communication – for example creating a profile on Facebook/Instagram/Snapchat;
- State – for example census data, CRB checks, electoral roll;
- Technology and services – for example internet browsing history, signing up to a mailing list, smartphone apps tracking location.

While these different situations came with different expectations of privacy, not all data fell neatly into one of these categories. In the subsequent discussion it became clear that new ways of collecting and sharing data, under new circumstances, can give rise to conflicting expectations around data privacy.

For the public each type of data-sharing activity is seen as naturally falling into a broader set or **type of transaction**. We have identified four different types, each with a different set of 'rules' and assumptions that govern the exchanges and interactions of that type.

Consumer transaction types (see Figure 1 below). In these contexts, the public expect lots of data-sharing, and are happy for this to happen because it is part of the 'deal' of buying something.

- BUYING: A customer allows a supermarket to collect data about their purchasing habits and brand preferences in return for the rewards of using a loyalty scheme; or they fill out a form online in order to get access to a cheaper car insurance policy. Data is **actively** given and treated as a sort of commodity in these transactions, with people receiving a clear personal gain in the form of a product or service, and with consent and in full knowledge that the information they give forms part of the transaction as well as the financial cost.

> *"The store card one is a relevant one – you get money off for things you may have bought and may buy again."*
> General public, Swansea

- DOING: This is a newer type of transaction, characterised by **passive** data collection. This might involve a smartphone collecting data about where you travel, or how many steps you take, or which apps you use, and passing those on to other companies for further analysis. It might involve social media targeting you with segmented advertising, while you view your profile. There is not a specific gain for the user associated with each data collection experience, but the most savvy users understand that they are generating data for the company; and that if they want the benefits of the technology, this is the bargain on which it operates. Less aware users, however, do not think about this and are potentially more vulnerable to exploitation.

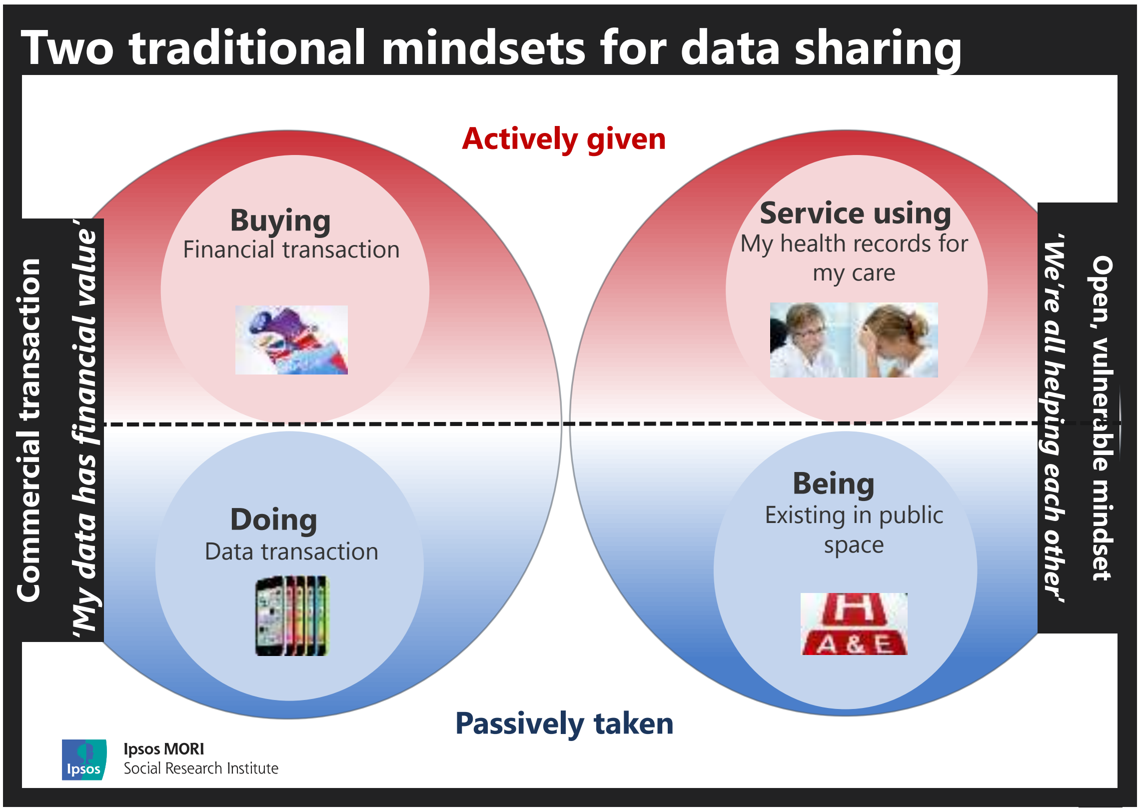
In both of these situations, transactions are usually characterised by a **wary and careful** approach. The individual believes that the company will pass on and sell any data it can, and he or she takes the responsibility not to be 'ripped off'.

Social Contract experiences. These experiences take place with the individual in an unguarded state; offering information they have no choice but to give. Participants bring an open, vulnerable mindset to these experiences, expecting help and support and feeling they are protected by a social contract in which we all contribute to services and they are run in our interests by benign authority figures.

- SERVICE USING: When attending a GP appointment, reporting a crime, visiting A&E, or the Job Centre, members of the public will **actively** disclose whatever they need to, as they are applying for help to live their lives. There is an expectation that individuals are protected by law, and data will not be shared without explicit consent. Participants felt strongly that this should be the case, because people in these situations are often suffering misfortune, which makes them vulnerable.

- BEING: A similar mindset is in place when people are relaxing at the park, or using public transport, or even entering and leaving hospitals or schools; just 'being'. In these contexts, expectations are of limited sharing, in order to protect privacy rights. There is traditionally an expectation of peaceful anonymity in public spaces, as individuals are not able to be identified and can move without hindrance. Participants think of data in these settings as **passively** given. As with the passive data collection in the commercial mindset, though, the public tend to underestimate the amount that is collected. They can only imagine light touch, aggregated statistical analysis. For example they suggest that the electoral roll or census might collect the number of people living in an area or sending children to school, while a hospital might count how many people went through A&E.

Figure 3.1 – The two traditional mindsets for data sharing



The specific context of a data transaction affects: the mindset the public feel they need to bring to the transaction; their expectation of privacy; their expectation of personal or social gain; and their perceived level of vulnerability to exploitation.

However, **concepts of privacy are in flux**, and participants in this research are aware of this. The **online world sets expectations of contexts which sometimes do not map onto the old contexts we know from the offline world.** Many people have not yet caught up or are not able to make sense of the increasing overlap that exists between the different transaction contexts.

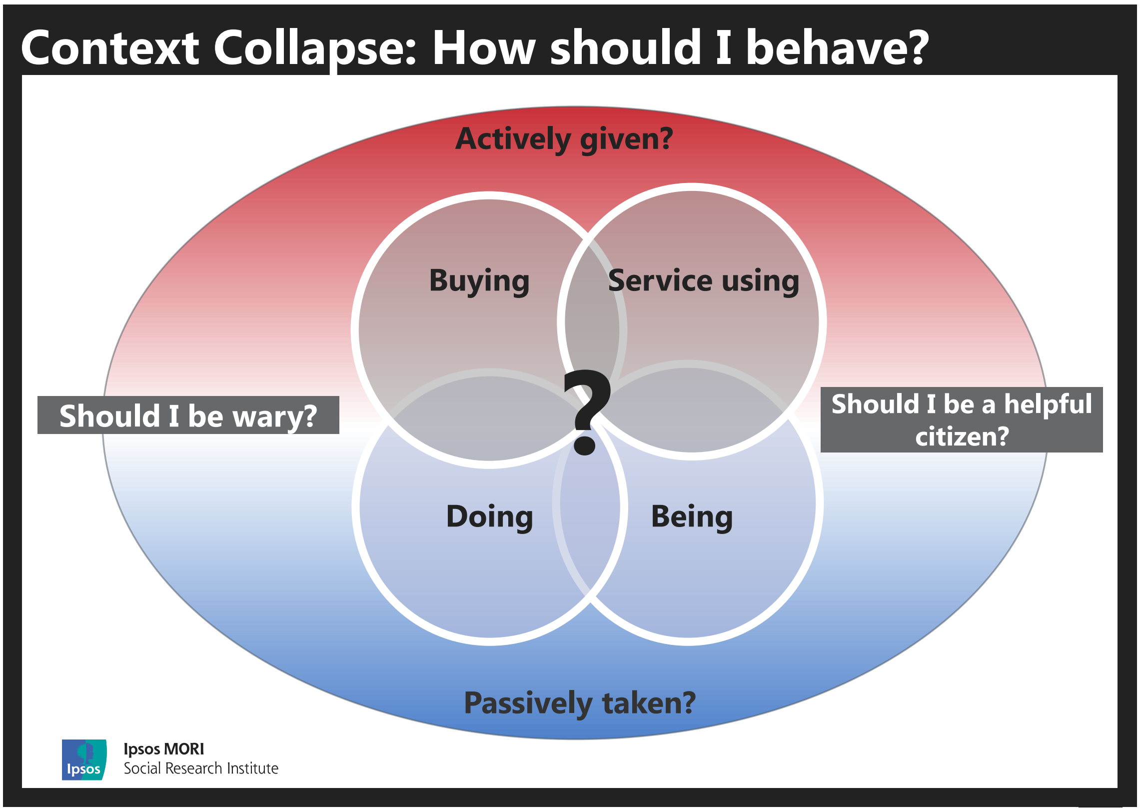
Passive data collection in both contexts tends to be underestimated; there is no explicit entering of the data and no consent at the moment of collection. Participants do not know how much is collected or how it is repurposed, and seriously underestimate the extent to which all data is analysed and looked at together with other datasets.

Commercial access to health data represents a '**context collapse.'**[10] This is illustrated in figure 3.2 below.

---

[10] A concept invented and disseminated by danah boyd, this paper is the first recorded occurrence http://www.tiara.org/blog/wp-content/uploads/2010/07/marwick_boyd_twitter_nms.pdf

Figure 3.2 – The 'context collapse'



Commercial organisations having access to health data cannot be interpreted as either a solely consumer or solely public data transaction. Information is shared in one context -when people are in a state of 'service using' or 'being' and of a vulnerable mindset – and is sometimes re-used in a commercial, transactional context. This means that the appropriate mindset for the public sector; open, vulnerable and helpful; may not be appropriate for the private sector, where a more guarded approach is called for.

Even if there are public benefits to the work that the commercial organisation does using the data, its presence in the transaction creates a contested and contradictory set of expectations and causes anxiety among the majority of participants. Participants did identify some opportunities within this changing situation, but also feared that it could cause harm to all, and especially to the most vulnerable.

They do not currently feel in a position to evaluate or accept this new data-sharing scenario. **Therefore, they fall back into their assumptions and personal beliefs and prejudices**. They look for lots of reassurances in order to feel more secure.

"Information about your health should be kept in the hospital, shops should be kept with the shops."
General public, Sutton Coldfield

Here are three examples of some discussions participants had, and how they illustrate underlying concerns about some of the changing 'rules' and context collapse.

This is all supported by the findings of our digital anthropology study. Key ideas included the notion that privacy is performed in contextual frames[11]; and that privacy is about control over accessibility of 'the private'[12]

1) Privacy is changing; we no longer have the 'expectation of uninterest', that people will **not want to overhear us in a public space**

Traditionally, when chatting to someone on a park bench, we don't expect to be listened to; there is a reasonable expectation of privacy in a public space. However when chatting to someone online, it is a public space 'with a difference' - it is possible for a third party to keep a record of everything that we do. Participants were concerned that an **unlimited amount of data was collected online** without our permission or knowledge.

"If I go on a website and look something up, all of a sudden other people are picking up ***this information. If you look at a PPI site, all of a sudden you're being*** *shared."*
General public, Glasgow

"We live in this age of technology but we have this choice, we are not numbers, ***we're people, we should have the choice to share our data or*** *not."*
General public, London

Participants showed little or no understanding that **content they posted on social media** could form part of their 'data profile', and could contain important information about them, beyond the immediate transaction of creating an account.

They mentioned individuals actively putting a lot of information about themselves on their profile for others to see. They knew that Facebook makes money from advertising to its users. However, in general people **did not discuss or seem aware of Facebook (or others)** monitoring and analysing the content of its users' posts to learn more about them. This was something that did not come up spontaneously in any workshop. This illustrates that they are concerned about the possible open-endedness of data sharing, but many are unaware of the ways in which it operates currently.

2) A meeting with the doctor is seen as 'service space', organised by the public sector, **where we can be vulnerable; not somewhere we have to be on our guard against commercial exploitation**

When talking to the doctor or pharmacist, we have a set of assumptions; we expect our information to be shared for our care but not to go beyond this and be used outside of that direct patient care context. When learning of commercial access to health data and current practice participants become confused and uneasy, are unsure how far they are able to veto

---

[11] Boyd, D. (2014). *It's complicated : the social lives of networked teens.*
[12] Nippert Eng, Christina E. 2010. "Secrets and Secrecy." In *Islands of Privacy* , 21–96. University of Chicago Press.

or give permission for any sharing to happen, and do not know whether they are putting themselves at risk or not by sharing.

***“I’m all up f*or improving and helping out research – *but I don’t think it’s right for them to just presume you’re ok with it and forward (your medical records) on* without your consent."*
General Public, Wrexham

***“Any sort of illness you have they’ll know, if you had* any sort of mental illnesses like depression, do you want your employer to know that?”*
General public, Sutton Coldfield

3) Sharing data in consumer space leads to ‘bombardment’ of marketing; people who don’t know this often adopt the wrong mindset and can **be vulnerable to exploitation**

Common to most groups was an irritation about continual marketing from companies, especially from those to whom participants believed they had not volunteered their contact details, or did so unwittingly. This was a concern about third party access which will be discussed in Chapter 4.

***"They shouldn’t be able to pass on your information, and they do. You get random* calls from people who have your landline number."*
General Public, Sheffield

*“When you get calls from companies – I get a lot of those – **you think ‘how did you get my number?’ Obviously it’s being shared without my knowledge, but** (originally) I shared that.”*
General Public, London

Participants felt that if they engaged with consumer life at all; opening accounts, using their phone, ordering things online, subscribing to mailing lists; this meant that their contact details would inevitably be passed around companies, who would then contact them. This was seen as a major annoyance of modern life and many mentioned that they were wary of all unsolicited email, post, texts or calls.

Participants felt uneasy about this on their own account and this manifested as some concern for elderly or vulnerable people, for whom a personal contact would likely trigger the ‘open, unguarded mindset’ not the ‘cautious, wary consumer’ mindset. Participants thought they might be open to **fraudsters** or unscrupulous marketing.

In fact, younger people may be equally at risk given that they do not necessarily know the full extent of the commercial data sharing they undertake online.

We heard less about the reverse idea, that data sharing might make you **wary, in a context in which you should be open.** There were some indications that this might happen; for example when discussing pharmacies, participants were concerned that trust in pharmacies might be eroded if data is shared with retail pharmacies, and therefore people would feel less comfortable going there for advice.

4) The slippery slope – a concern that data given and shared in the private sector for a transaction might be used by the public sector and might affect my rights and services

For some, the idea of information being shared without their knowledge led to fears of a dystopian society where all of the information they have ever provided – for example on an official form or when interacting with the Police or the NHS, plus store card or online data is held in a centralised system and linked up.

While there is little sense of how this might realistically be done, who would do it, and what they would want it for, this tends to create fears that privacy is fatally eroded.

4 The factors influencing attitudes to commercial data sharing

4 The factors influencing attitudes to commercial data sharing

4.1 Using case studies to explore attitudes to variables

During the qualitative workshops, participants were introduced to **six case studies** (or four, in the shorter workshops), as examples of real-life scenarios in which the sharing of health data currently takes place (see Appendix for stimulus materials)[13]. The case studies gave examples of existing data-sharing activity in the public and private sectors and were followed up with discussion of hypothetical future activities (including deliberately provocative and controversial ones).

The aim of the case studies was to show participants the range of different types of data sharing with commercial organisations in order to elicit the boundaries and principles which were important.

Small groups of participants were presented with the case studies and the facilitator discussed each in turn. In discussing these case studies, participants were encouraged to weigh up how they perceived the benefits to themselves and society of these different examples when considering acceptability. They thought about the risks and potential for harms, both for individuals and society. They tried to imagine what future risks or benefits there might be (relevant particularly where such future risks and benefits are uncertain, such as in emergent areas of data science or when considering genetic data).

---

[13] The case study 'Monitoring safety of drugs and medicine' was first introduced to participants as being conducted by a public health regulator. This was intended to contrast with the five other examples of commercial access so that the research could explore how responses might be different if a non-commercial organisation was involved. In follow-up discussions, the regulator was replaced with a pharmaceutical company and the question of commercial involvement was then further explored.

Summary of the six case studies

- **Data linking and analysis in the NHS** – An NHS trust asks a healthcare intelligence company to analyse individual-level data on patient journeys, to see if there are different patterns in health outcomes, and predict drivers of service use.
- **Monitoring safety of drugs and medicines** – A pharmaceutical company runs an observational study to look at long term side effects of a blood pressure drug. Primary care data is provided to compare the probability of serious adverse events among those taking the drug compared to those on other drugs for high blood pressure.
- **Calculating insurance premiums** – Private health insurance companies use anonymised hospital data about diagnoses and hospital admissions and find that those living in deprived areas were more likely to develop certain critical illnesses.
- **Using genetic data in care and research** – Patients consent to having their genome sequenced as part of their clinical care. This is linked to their medical records to aid diagnosis and treatment, and made available for research by academics, scientists and commercial organisations.
- **Pharmacists using Summary Care Records** – The NHS wants all community pharmacists to have access to a summary care record. Pharmacists would have access to this with patient consent when discussing prescriptions.
- **Crowdsourcing to provide support for patients** – Patients register on a free online community to share experiences and symptoms. The online community allow a drug company to invite diabetics to participate in research into the efficacy of a drug to treat sight loss.

Once they had discussed all of the case studies in turn, participants then **mapped them according to their value to society versus their risk to society.** This exercise required some level of consensus to be reached in the group.

This brought out more nuance in their views, as they moved away from focusing on the more easily recognisable and personally relevant considerations and focused more broadly on implications for society as a whole.

Figure 4.1 – Participants mapping the case studies

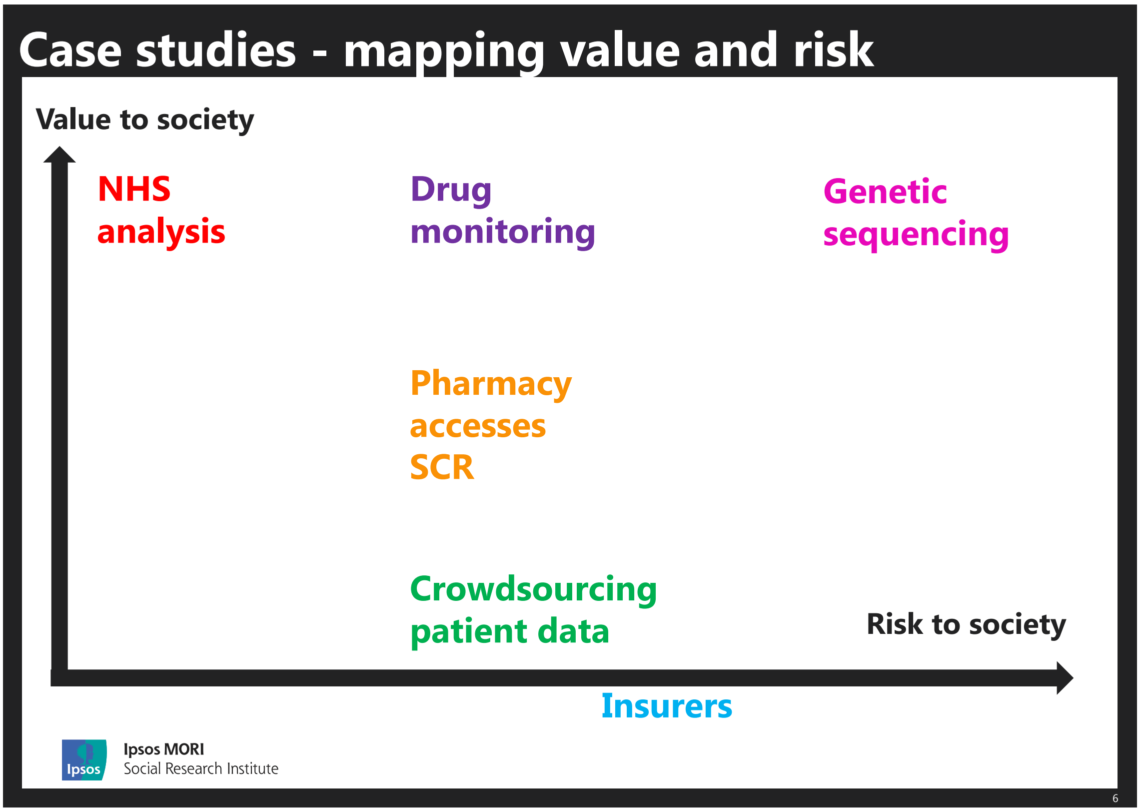
4.2 What mattered overall: response to the case studies.

4.2.1 Some examples were seen to be of higher value, or higher risk, to society

We asked participants to map the examples according to whether the commercial access would bring ‘value’. However, we did not provide a definition of value; keeping it open enabled us to understand how participants themselves conceptualised value in this context.

Some commercial access types were seen to have a higher value, or higher risk, to society than others:

Figure 4.2 – Mapping value and risk



Case studies - mapping value and risk

The social benefit of commercial access is not always apparent at the start of the deliberation process so people either focused on risks to themselves or exaggerated the public risk and feared the worst of private sector involvement.

Without a clear conception of how the public stand to gain, the public cannot carry out their internal trade-off exercise and weigh up public good against personal risk. They often revert to their pre-existing stereotypes about government and business in the absence of more knowledge.

This all spotlights the role that *understanding the new context of health data-sharing* plays in public attitudes towards commercial access and the role of raising public awareness and engaging the public further in discussion.

4.2.2 Different types of organisation and project were perceived differently

The involvement of a commercial organisation was seen as fairly easy to accept when participants could see **clear potential for patients, society and future generations to benefit.** This meant **Linking data in the NHS** and **Monitoring the safety of drugs** were acceptable and seen as valuable. Participants asked why the NHS "could not do these jobs itself"; suggesting that they can **accept a company being involved**, as long as it is for public benefit, and only if they are given information as to why companies can do this better than the NHS.

***"There's a certain level of accountability if it's being done by the government, but if it's a private company that is only accountable to its shareholders' profit, then** the public good is undermined."*
Patients (rare conditions), Sheffield

4.2.3 Participants were keen that public benefit should be maintained even if a company, rather than the NHS is doing the analysis

Participants had the idea that in **Monitoring the Safety of Drugs**, they could ensure that the private sector works to public interest rather than in its own interest by the private sector funding academics or a regulator to conduct the work. Some saw it as risky if the work is done by a for-profit organisation; this would increase the chance of bias. However, **if the** organisation can effectively 'prove' that it is working in the public interest, for example by **submitting to voluntary regulation or oversight,** then there is trust that the organisation could be allowed access to health data.

"The pharma company should pay for it, the regulator or academics should do it."
General Public, Sutton Coldfield

4.2.4 There are concerns about large commercial organisations such as supermarkets assuming a role that they felt should belong with the GP and the NHS

The pharmacies accessing patients' **Summary Care Records** example was considered to be potentially a slippery slope, where the role of the NHS might be privatised and eroded. Participants questioned how far these organisations could and should be trusted with public benefits, when they had conflicting commercial priorities.

However overall this example was felt to be quite valuable to society, in that it could potentially make the prescriptions system more efficient, freeing up GPs' time as well as making it easier and quicker for patients to receive the medications they need. **Anonymisation** was impossible to preserve in the case of the **Summary Care Records** and therefore the **opt-out** (which is part of the process) was felt to be particularly important.

The example of **crowdsourced patient data** being used for research was seen as being of fairly low value, and moderate risk;

Although participants appreciated the potential for this to contribute to the development of new drugs and treatments, there were concerns about the **accuracy of the data involved, and the vulnerability of the system** to misuse and hacking. It seems the public will need evidence of the value of crowdsourced data in order to trust that a project including such data can have a true public benefit.

4.2.5 Genetic sequencing was considered to be the most risky example; genetic data both most private, and most potentially valuable

Participants were concerned, overall, because so much is currently unknown and yet to be discovered in the field of genetics and thus the full extent of what might be possible with this type of data is also an unknown. This made it harder for them to weigh up benefits and risks and meant that most were cautious and felt it might be both risky and valuable.

They wanted **universities and independent researchers** to be bodies who would, ideally, have access to it. Any link to non-healthcare companies, such as marketing and insurance, was totally unacceptable.

"I'm more than happy for academics or researchers to see this, but not private companies."
General Public, Sheffield

"They sound like private sector, and what might they use the data for? NHS clinicians need to know, academic researchers would have a positive effect, but genomic technology companies – what do they do?"
General public, London

4.3 Participants applied four ‘key tests’ to every example of data access

4.3.1 The four tests

Participants discussed the issues arising from each case study as they saw it, and came to conclusions about how acceptable each might be. Their approach could be summarised as applying **four key tests** to all the case studies. These tests must each be passed before the public will accept commercial access to health, biomedical or genetic data in that example.

1. WHY (Does the activity’s outcome have a provable and sufficient public benefit?)

2. WHO (Can the organisations doing this be trusted to have public interest at heart?)

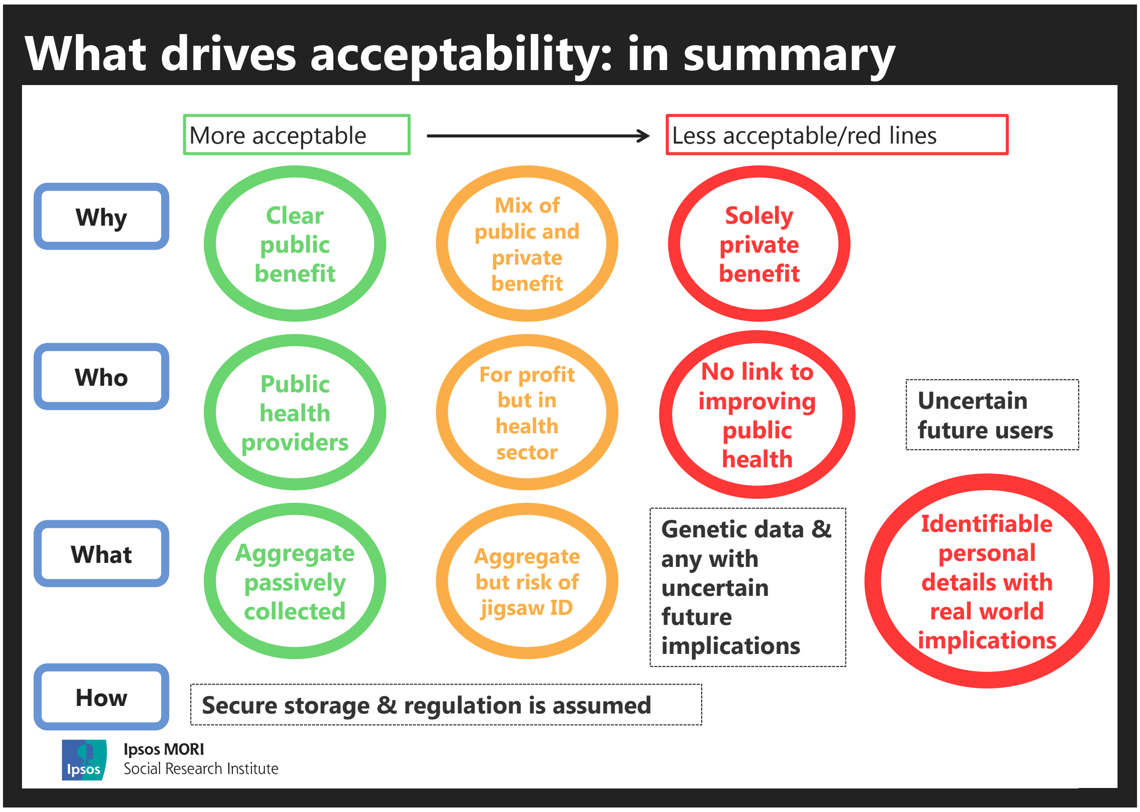
3. WHAT (How anonymised and aggregated is the data?)

4. HOW (Does the safeguarding, access and storage protocol reassure me that the data will be safe?)

The Wellcome Trust had hypothesised that these factors might be important on commissioning the study. The deliberative work confirms this and crucially discovers that the tests are always applied in the same order. Failure to pass one means that the case study is immediately ‘thrown out’. So, for example,a case study with no perceived public benefit at test 1, but ‘scoring highly’ on test 3, (for example using very aggregated data), was not enough for participants to feel happy about a commercial organisation being given access to the data.

Figure 4.3 shows levels of acceptability for each test question and the types of organisation and data which are envisaged at each stage. The red colour suggests the data sharing situations which are broadly unacceptable; amber indicates situations which depended heavily on context and circumstances, or about which participants were unsure or divided in opinion; green denotes situations with which the majority of participants were comfortable. For each, there were some exceptions, as discussed in the next chapter on Mindsets.

Figure 4.3 – The four 'key tests' and factors which drive acceptability of commercial access to health data



What drives acceptability: in summary

More acceptable → Less acceptable/red lines

Why
Clear public benefit
Mix of public and private benefit
Solely private benefit

Who
Public health providers
For profit but in health sector
No link to improving public health
Uncertain future users

What
Aggregate passively collected
Aggregate but risk of jigsaw ID
Genetic data & any with uncertain future implications
Identifiable personal details with real world implications

How
Secure storage & regulation is assumed

Ipsos MORI
Social Research Institute

This process, and the order of importance of each test, remained the same for all groups within the workshops.

This resonates with the key findings of the **quantitative** work. In the survey, there was overall support for health data being accessed by commercial organisations undertaking **health** research (54 per cent supporting, compared with 26 per cent opposing). However support for sharing health data for the purposes of **insurance** or **marketing** was lower. Just a quarter (26 per cent) support sharing anonymised health records with insurance companies so they can develop their insurance prices. Support for companies marketing health products using anonymised health records is slightly higher, with 37 per cent supporting this purpose. For more detail, see chapter 6, where the full findings from the survey are presented.

4.3.2 WHY: Does the activity's outcome lead to public benefit? The most important factor

This project was designed (in part) because surveys suggest that the public are concerned about commercial access to health data. As it turns out, this is something of a sweeping statement. In fact, views are very dependent on the nature of the organisation and its purposes. For most, if the overall purpose of the data-sharing activity is considered acceptable, concerns relating to the commercial nature of the organisation(s) involved often fade.

Participants in this research assumed that privacy should be the default and that any sharing would infringe that privacy. There was no assumption that data should be automatically in the public domain. A clear **benefit both to individuals and to wider society** was seen as the only good rationale for breaking privacy, and this was the primary driver of acceptability for participants.

Conceptions of "public benefit" were very broad and encompassed both benefits to individuals or groups in society, and benefits to big public institutions like the NHS. Participants tended not to use the specific term "public benefit" but instead used phrases like "if it helps us", "if it is in our interests", or asked "what does the NHS get out of this?" to describe the concept of a public return for data-sharing. Examples of public benefit participants liked included:

- Developing a life-saving drug;
- Allowing patients to collect repeat prescriptions from the pharmacy instead of the GP;
- Improvements in paediatric care nationwide;
- More sophisticated or successful treatments and diagnosis for illnesses;
- Higher levels of service, correcting mistakes, or creating greater access to services for vulnerable groups.

> *"If you can share information about conditions, more research can be done into* ***that condition, it can help someone else (…) More medicines can be produced,*** *people can be cross-checked and they can find solutions to genetic problems that people have in their system."*
> General public, Belfast

Individual citizens gaining benefits was also seen as valuable, **but would have less value than all of society benefiting**, which is why the Pharmacy and Crowdsourcing examples were valued less highly than the NHS analysis, Drug Monitoring and Genetic Sequencing.

Red Lines: no public benefit

When no benefit to public health is perceived, commercial access is unacceptable to all, except the most laissez-faire (see Chapter 5 for description of these mindsets).

4.3.3 WHO is doing this, and whether they can be trusted to have public interest at heart, is almost as important

Overall, participants paid close attention to the *type* of organisation involved as this helped them come to a judgement about the purpose and likely outcome of the activity. They were notably more comfortable with some handling health data than others. There was a strong desire for data in general to be kept within the healthcare sector.

Whatever the organisation, participants in the qualitative survey as well as many in the quantitative survey felt strongly that the onus lies with the commercial companies to demonstrate that they are creating social goods within the realm of healthcare and medical research. In the quantitative findings, nearly half (43 per cent) wanted commercial organisations to show 'a clear intent that research will lead to benefits for wider society'.

The 'WHO' can affect the 'WHY', because if the public do not believe the organisation capable of achieving disinterested public benefit, then the first test has been failed. However, there are ways even the least publicly-beneficial organisation can pass the test, for example as mentioned previously (section 4.2), **organisations can subject themselves to regulatory scrutiny or partner with academics or charities to enhance their reputation**; thus ensuring that the outcome is *provably* beneficial.

HCPs and general public participants are most open to the idea of academic researchers, charities, or partnerships between these and the public sector having access to health data.

Most were not used to thinking about medical research as a space in which public and commercial organisations work in partnership and initial reactions were often of surprise and sometimes shock. **Those with greater awareness and understanding tended to be more** open to the idea of commercial organisations' involvement.

There is further evidence for this in the quantitative research, where those with a good awareness of commercial organisations having access to health data can be seen to be more likely to support commercial organisations having access to health data *for research purposes* (see chapter 6).

GPs and hospital doctors felt that having researchers work in partnership with commercial organisations provided a very useful **safeguard** against unethical conduct.

"I do work in a university - ***partnerships are okay. I'm not sure whe****ther drug* ***companies are particularly good (...) but that's why teaming up with third sector*** *or universities gives a bit more robustness."*
GPs, London

Other healthcare professionals were less worried about private companies' involvement and saw similar potential for misconduct where research is carried out by public and third sectors alone. They pointed to poorly stored data and inadequate security checks and an informality in how personal health data is shared, emphasising that the public health sector should treat data more responsibly.

"[I worry about] the accuracy of what is shared, before you know it there is chaos! The problem is you almost want someone who knows about diabetes overseeing things."
GPs, London

There was even support among GPs for **more commercial involvement**, as they believed that in certain scenarios they were better placed to use health data objectively.

"There is some element of fairness (because) if it's the NHS (measuring) itself, it's dodgy dealings."
GPs, London

Patients had the greatest understanding of the interplay between different types of organisations in medical research (e.g. charities, universities, sponsors and commercial organisations). In particular, some of those with rare diseases had personal experience of being involved in research projects and were familiar with the important research work of charities. They firmly believed that charities should do *more* with data (and that regulation should enable, not prohibit). Among the general public, there was a little more scepticism about charities.

These non-commercial organisations were allowed to access health data because participants strongly associated them with a clear public benefit. It was assumed that they would use the data responsibly and with public interest in mind, considering both how to *protect* individual privacy and safety and how to ensure there is a *public benefit* outcome.

Beyond researchers and charities, there was a **hierarchy of acceptable commercial organisations:**

- Specialist analytics and research companies working closely with the NHS; participants did not know much about them, but felt they were required to provide public sector value in order to make money, so were inherently benign.

- Pharmaceutical companies; while some feared 'Big Pharma' had an agenda at odds with the public interest, most acknowledged that a regulated pharma sector helped public healthcare and needed data to do its job. Regulation was required, however.

- Retail and pharmacy sectors; many participants were uncomfortable with the reality that a retailer or manufacturer's ultimate objective was to drive profit for shareholders rather than create public value.

"Big pharma ...Are they doing it with my consent, looking at a group to identify, make progress, come up with treatments, understand conditions more – I'd be comfortable with that. Or are they just given free rein on my daughter's medical records so they can stabilize business, play entrepreneurs, gamble on it - no, that's not ok."
Patients (rare conditions), Sheffield

"Why would there be a company [involved], I guess it will make a profit at some level."
General public, London

The quantitative research found that though support for marketing and insurance purposes was lacking, acceptability of pharmaceutical companies conducting research using health data did not differ significantly from the acceptability of a public health regulator conducting the same research. This was identified using a split sample experiment that is further outlined in chapter 6. These findings suggest that though there are concerns about different commercial organisations having access to health data, acceptability is dependent on the type of organisation.

While the quantitative work does not explore why this is, our experience with the qualitative work suggests that pharmaceutical companies, being more accepted as organisations working within the health system, are perhaps more readily accepted for using health data than organisations without a healthcare core purpose.

Where the organisation's purpose was **uncertain**, or its purpose unclear (for example a proposed future genetics data company) participants mostly erred on the side of caution and assumed that the profit motive would be present, except for the most laissez-faire (see Chapter 5 for a discussion of this group).

Red lines: insurance

Reactions to the idea of an **insurance company** using health data to adjust premiums was met with universal disapproval. This dislike was also seen in the survey findings, which show that just a quarter (26 per cent) accept sharing health data for the purposes of developing health insurance prices.

Participants could not see any public benefit to this, and viewed insurance companies with mistrust. It was felt that the data would only ever be used to increase premiums and make life more difficult for people with health conditions, and there were worries about the data being traced back to the individual. Participants said they always begrudged giving insurance companies access to their personal records, as the industry is seen as having high charges but often refusing to pay out; the public believe this would get worse if more health data was available to the industry.

There was a more subtle objection, also. Many participants saw the motives of the whole industry as entirely at odds with public interest. Some said health insurance companies were **contradictory to the basic principle of a public health service.** Health insurance of all kinds was also seen as something that only affluent people would choose to have; the industry was perceived to be harmful to **society**, working against the principle of a public health service.

In Sutton Coldfield, the improvising actors who were watching the session exaggerated some of the things they heard for comic effect. One of the most popular sketches they then improvised was about an insurance company putting prices up again and again after an initial offer of a very low price; this resonated with the participants.

"Have you ever had a grazed knee? …At any time in the last forty years? Yes? ***Well I'm afraid your premium has*** *now gone up."*
Comedy sketch by improvisers, Sutton Coldfield

Red lines: marketing

Further discussions of a similar role for **marketing companies** led to more nuanced responses. Some participants recognised that targeted marketing was a fact of modern life and could, in some cases, be beneficial to them personally. Others, usually older participants, voiced strong dislike at what they felt were often aggressive and intrusive direct marketing techniques.

When the marketing is for healthcare products, there was a sense in the workshops that this might be more acceptable, that it might just serve the public interest in some ways. In the quantitative survey, as many agreed as disagreed with sharing health data for marketing purposes. While this agreement is higher than insurance companies it is still a lower level of support than the survey found for commercial access in general. The example of marketing purposes that was shown to respondents was based on targeting an area for a health product (a low-fat margarine) and so it might be one of the more acceptable scenarios that could be used.

Marketing companies were usually associated with unwanted telephone calls and junk mail which carry negative connotations and are the source of much public frustration.

Some attitudinal groups in the qualitative work also felt more relaxed about marketing; which we discuss in Chapter 5 on mindsets.

Red lines: third party access

Not allowing the data to be sold on to other organisations beyond its original use was universally important; participants said on many occasions that they did not want data passed on several times or re-used. These conditions helped to make **Linking Data in the NHS** and **Monitoring the Safety of Drugs** acceptable.

For some, this applied not just to data, but also to any re-analysis based on the original data.

The more steps are taken to repurpose data from its original collection purpose, the more worrying this feels for people. Even if the reason for third party access was deemed important, if participants did not trust the organisation conducting the research they called into question the ultimate public benefit of the research.

In particular, it seemed to go against natural justice that a company could **repurpose data originally generated from the public, and make money again and again** from the same dataset.

***"The decision is clearer (when the data is being used) by a public regulator, it's** not a private company and it is being used for the purpose that it was collected for."*
Hospital doctors, Birmingham

Red lines: a need for a regulatory safeguard

Participants expressed fears of companies 'spinning' the social worthiness of their activities in order to gain access to data; there is a real role for regulation and governance.

4.3.4 WHAT: How anonymised and aggregated is the data? Matters more to some than others

Participants accepted the definitions given of **aggregate data**. Where they heard the word 'statistsics' and understood that these were only numbers taken from the aggregate data, participants did not feel this type of data was very risky to the individual. Some, however thought that this data even in aggregate could be **risky to society** if re-purposed in ways that led to groups being discriminated against.

However, **all other kinds of data were considered a potential threat**. People were concerned about the theoretical risk of jigsaw identification from anonymised individual level data and assumed that if it was possible, it was a risk. Individual level data was seen as private. Concerns tended to arise when the aggregate data was made up of large numbers of individual patient records, which had to be stripped of identifiers, and *this* was shared with commercial companies.

Healthcare professionals and Cohort members were more **relaxed about allowing commercial access** to aggregated data because of their greater familiarity with the process and understanding of the lack of risk.

***"When it's at a personal level you need the limitations. But when it's on a general level and they don't know who you are, it's not suc**h a big issue."*
Cohort members, Bristol

Some participants (see Chapter 5) felt that we are heading for a dystopian, surveillance-based society. For them, whether the data is anonymised or not was almost not relevant because they were worried about the future risks to privacy rights as a result of so much increased data sharing **of any kind**.

Many participants saw commercial access to **genetic** data as a red line because the exact content of any data that might be shared in future is currently unclear.

A long term and far-reaching discussion was felt to be needed around this subject, including government, citizens, the research and academic community, lawyers and business, before conclusions could be firmly drawn as to what should happen to our genetic data. At the start of such a process, participants could only be cautious and so most tended to prefer access to be denied, pending this debate.

"This is sort of future stuff. This is a problem for maybe 20 years' time. In 20 years' time this might affect you."
General public, London

Red Lines: personal harm

For example, where **demographic data** is crossed with health records and used to map the country according to the likelihood of local populations to develop a critical illness. Participants worried about implications for people living in areas designated 'high likelihood' and the potential for it to get into the wrong hands and impact negatively on **vulnerable individuals**, with already vulnerable members of society being pigeon-holed and penalised.

"If you're poor you'll get poorer. You'll remain in the area as you can't get out of it."
General public, London

4.3.5 HOW (Does the access and storage protocol reassure me that the data will be safe?) Mainly only an issue when people see personal risk

Most participants believed that in today's world of health data-sharing, good data security and protection is a basic **hygiene factor**. They often **assumed that security systems would be in place** that were appropriate to the nature of the activity, data and organisations involved and that only skilled professionals would ever have direct access.

In general, participants tended to say that the data would be safer in the hands of the NHS or a public sector or independent organisation, and that private companies were less likely to be as diligent in their handling of it. Thus NHS involvement acted as a reassurance that the data would be kept safe.

Attitudes to secure storage and handling were largely affected by participants' general understanding of **key technological and security concepts**. Many were used to thinking of security in non-digital terms and used this as their frame of reference, wanting assurance that data will be "stored" in a safe place and "moved" safely between individuals and organisations. Older and less tech-savvy types used this kind of terminology, reflecting the confused sense they have of how data is used.

Cloud databases caused worry among those with little experience of using them and participants feared the seemingly open aspect of a cloud; they assumed more parties would be able to gain access than originally intended. Some conflated *a* cloud database with *the* cloud database and naturally became concerned for the safety and accessibility of the data.

"I never use the cloud, the cloud is not safe, my kids can get on it."
General public, Swansea.

This illustrates the challenge in setting up any safeguards which will reassure the public. Because people did not really understand what datasets are, or how they are used and stored, there was little understanding of why safeguards would work, and hence what safeguards would be best.

There was **little understanding of the status quo** (many of the safeguards workshop participants asked for are actually in place already).

Participants believed that **data has a unique nature**. If the potential exists to re-identify, at any level, within the data, despite storage and safeguards, they want to proceed cautiously even though in practice the risk might be minimal. There was a real assumption that mistakes inevitably happen (especially for the workshops taking place in the week that TalkTalk lost a lot of customer data) and no assurances could really alter that state of affairs. Also, some believed that **future governments could change policy on this**, and if the data has been collected, it might be used in future, whatever we say now.

Participants believed that no amount of security could ever totally remove the risks involved in sharing data such as **leaks and hacking**. Plus, however secure a dataset may be now, participants believed that there could always be **unknown future risks** that cannot be controlled for. The introduction of commercial organisations exacerbated this worry as it was seen by participants as an indication that publicly held data is moving further away from the public systems accountable for it. This connects to fear of loss of control.

"Who knows what will happen in the future? Anyone could get hold of this data."
General public, Sheffield

Red lines: communal, open access to identifiable health data

In the pharmacy case study where this data is used to provide personalised direct care to patients, participants were clear that they wanted to limit the number of individuals with access to their personal information. There were fears of open plan shops, which might allow an opportunity for others to read or overhear personal records. Participants expected access to be limited to employees with professional medical experience.

4.4 The implications for public trust in safeguards

The public in the qualitative research felt that when given access to health data, **researchers** are contributing useful skills and knowledge to a modern healthcare system. This is also perceived to be the case for **regulators**, **charities** and some businesses, for example **social enterprises**. Beyond these organisations, participants felt businesses need to make themselves open to independent scrutiny.

"Licensing and regulation – if people are doing it for good then they should be happy for it to be regulated. Every organisation should be licensed and regulated ***and audited to ensure that there's that level of openness."***
Patients (rare conditions), Sheffield

The call for safeguards reflects a wider concern about governance of data and the need for a greater discussion of how to regulate the outcomes permitted to come from data sharing.

Participants asked for a range of things, not all of which are practical but which reflect their preoccupation with equity, transparency and a need to prioritise public benefit over profit. They wanted:

- The amount of profit that could be made from data given to the private sector to be capped; this would put the emphasis back on public benefits not profits. This illustrates the strength of feeling around preventing profiteering.

- Third party access removed. Access to be granted separately for each project, so that data sharing will be 'one stage' only; and data not to be transferred from one organisation to another under the same consent or permission. While people were concerned over lack of control and accountability if datasets were passed on, there was also real indignation that the data could create large profits in the private sector. This access restriction was mainly requested in order to limit profit-making opportunities, even if in practice this would be hard to implement.

- Independent scrutiny, whether by ethics committees or other bodies free from vested interests (to ensure that the social health benefits are in fact in the public interest).

- Transparency in sharing and publicising results; this would also underline the need for public benefit to be clear throughout.

The public also asked for **storage and access** safeguards such as sanctions (to prevent misuse and data loss). Participants in the qualitative research admitted that they knew little about how these more practical safeguards would actually work, or how data was stored or accessed. But the idea that regulations were in place felt comforting.

The idea of only allowing **named individuals** within an organisation to access health data has little traction. Participants pointed out that the individual themselves may not be trustworthy. Many are reassured by the suggestion that every time somebody has access they should be required to enter a log. The public like that if anything goes wrong it is easy to trace the user and that in turn acts as a driver of good practice and means it is possible to monitor and enact consequences for excessive use. This came from the qualitative work, but could be further explored quantitatively if it is an idea worth taking forward.

"That would make me feel a bit more comfortable because they would know, if for any reason the system had been abused, not that it would be but they would ***know… There'll be a shortlist of people who have accessed it, would be a*** *deterrent of abuse."*
General public, Belfast

This is borne out in the quantitative study. In the split sample experiment, knowing that a safeguard is in place, regardless of the type of safeguard, makes a difference to acceptability of health data sharing (agreement was 49 per cent for those respondents where no safeguard was mentioned; between 56 per cent and 64 per cent agree with data sharing with any of the named safeguards in place). But there are no statistically significant differences between acceptance of health data being shared with an **opt-out, anonymity, controlled access or robust governance** (in our example, heavy fines and prison sentences).

4.5 The implications for consent

The desire for the NHS to ask permission before this kind of data sharing takes place is widespread, with 54 per cent saying that they would rather the NHS ask patients' permission before sharing data with commercial organisations, even if this meant that without permission the research cannot take place. Younger respondents, those in the DE social grades and people in a BME ethnic group are among the most likely groups to see a priority on consent. The survey also found that a majority of people (53 per cent) want to see strict rules in place that data could not be passed to third parties.

The discussion of consent in the workshops reflected the way that participants did not necessarily know how large datasets are used and managed in healthcare.

The deliberative process may have made a difference to the discussion of consent in the workshops. Most participants in the qualitative work started from the premise that *everything* should be consented at all times, every time a piece of data was accessed by commercial organisations, and that access should not take place if this could not be guaranteed. However the majority gradually changed their views through the day to agree that this might be impractical. This suggests that more discussion on this subject may create a more evolved public view. This fits with the general mood from the survey that there is support for commercial access projects to go ahead *if* the benefits are seen to be worth it.

Because genetic data was sensitive the public assumed any data sharing would be opt-in, and this did not change through the day.

The surveyed groups who were concerned about consent seem impressionistically to be similar to the types of people in the qualitative research who focused more on harm to individuals rather than social harms or benefits.

It seems that concerns about consent play out differently in a survey than in deliberation; these younger people often had the first knee-jerk response that consent should be sought, but tended to change views when others started bringing in societal considerations.

When planning the level of consent which will create public trust, therefore, the different needs of different groups should be considered, and there may be a need to segment audiences more thoroughly when it comes to future engagement and communication.

Healthcare professionals felt consent became more important even with aggregate data if the private sector was to be involved. In particular, if third party access was allowed, this would make it important to get explicit consent, in their view.

Participants in the workshops talked about how the process of going about consenting could be improved, in the confusing new world of data provision and data access:

- Participants wanted a better understanding of **aggregation** and **anonymisation**. Clear and simple reassurance about how personal identifiers are removed from datasets and the unfeasibility of tracking down an average member of the public, will be important for building public trust.

- Participants felt that the world of **online** commercial transactions has created bad practice and '**confusion marketing'** around consent, with tiny boxes to tick, opt-ins and opt-outs where the default suddenly changes to trick people, and a conflation of agreeing to pass on details for transactional purposes with agreeing to marketing purposes. There is a need for a discussion about a better approach to consent. Participants asked for regulators, or future commercial data-sharers, to be held to high standards. They wanted data controllers to take advantage of the opportunities of online and digital settings to communicate more clearly, rather than the current situation, where the online world provides an opportunity to obfuscate and confuse.

- If it is not practical to seek consent for every piece of data sharing, participants agreed there needs to be more **communication around why this is**. Participants in the workshops, patients, and **GPs** too, felt that health care professionals could be trained to be good gatekeepers to explain how research works and what the implications are of giving or withholding consent.

In the sessions where **improvisers** brought to life the discussions, the actors listened to these debates and exaggerated them in the form of comic sketches. Some of the sketches picked up on the fact that participants feared the complexity of regulation in this new area; making the need for effective communication, to create consent, very important.

***"Hello! I'm from a community group, who scrutinises the private company who** watches the public sector who watches the private sector guy who watches the data to ensure all is secure."*

Comedy sketch from improvisers in Sutton Coldfield

Similarly, a sketch about very complex 'Terms and Conditions' with a baffled man trying to negotiate consent at his computer, created much laughter of recognition.

"I'm buying a new skateboard, and I've been sitting here reading the terms and conditions since last year. I'm hungry… I'm so hungry…"

Comedy sketch from improvisers in Wrexham

Overall, participants felt that if they knew more about the processes and safeguards in place they might feel more empowered, and hence more open and trusting in the decision-making process around data collection and sharing. In this instance, consent helps reassure you that you have control by opting out if you want to; if you knew more about it, however, you might not want to exercise the right to opt out.

> ***"****[Profit could be allowed if] p**eople are informed and consented and know it's** going to happen. Because then you know how to stop it if you need to.****"***
>
> General Public, Sutton Coldfield

5 Mindsets: what makes a difference to views?

5 Mindsets: what makes a difference to views?

5.1 Emerging response patterns in the deliberative work

In analysing the many different conversations held at the deliberative events, it became clear that patterns of response were emerging and people could be clustered into different attitudinal groups which explained their responses.

There are different reasons why individuals might hold an attitude. The non-linear relationship in the survey between age and acceptance of commercial access demonstrates that, while age is related to acceptance, there are other factors within the age groups that act as alternative drivers. Another example is when respondents who did not want commercial access to health data under any circumstances were asked why, respondents fall evenly into two camps: there are those who fear that commercial access to data might affect them or their family negatively, and those who fear that commercial access will have a negative impact on society.

These sub-groups within sub-groups are not possible to tease out in a short questionnaire, but the qualitative work allows us to identify differing attitudinal clusters that may cut *across* demographic variables. These are referred to as 'mindsets' and are a useful way of addressing the problem that there is no one clear concern or set of concerns that have been articulated. In fact, some people may hold some concerns based on certain worldviews, while others hold different concerns underpinned by separate beliefs about how society operates.

For these reasons, this chapter does not refer to the quantitative findings in relation to the mindsets. As well as demographics not adequately representing the complexity of the views within different sub-groups, the questions were not designed to address these queries.

This chapter describes the groups we observed and infers the principles which are important to each group by the ways they argued and the principles they drew on to make their points. It is important to note that the groups were not recruited to reflect any particular attitudes, but nevertheless each of these different mindsets was seen in different groups, to some extent.

The observed groups were partly associated with age, education, and sociodemographic status, but could more **accurately be described as mindsets.** Participants took different positions on risks and benefits and justified their views according to a variety of principles they espoused about society and commerce. The types of principles expressed differed between individuals, but formed relatively predictable clusters.

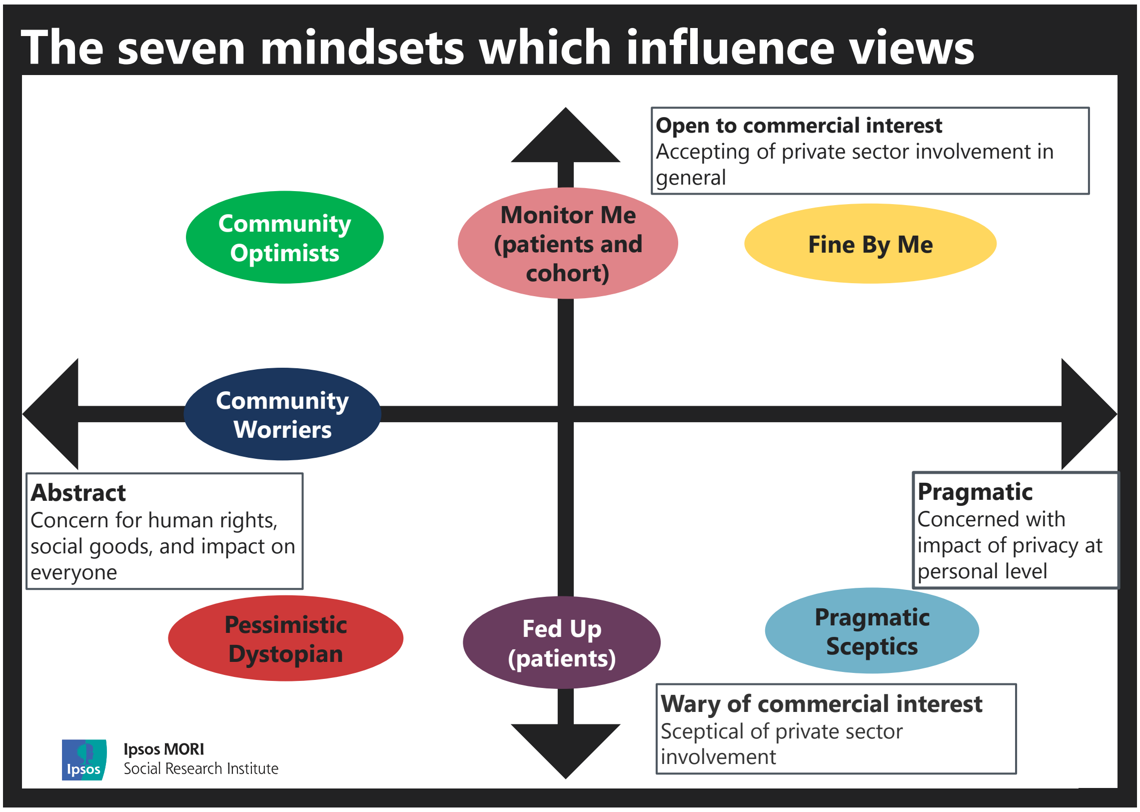
We spotted **five different mindsets** which shape the way the general public think about commercial access to health data, and **two mindsets more prevalent among patient groups**. Each mindset affected the way in which an individual reacted towards commercial access to health data and where they placed their **red lines of unacceptability**, when applying the tests of WHY, WHO, WHAT and HOW.

Views were broadly divided on two different dimensions:

1. **Open to commercial interest** and accepting of the fact there is private sector involvement in many areas of life VERSUS **wary of commercial interest** and wary of private sector involvement.
2. **Concern for pragmatic and personal issues** such as privacy and individual vulnerability VERSUS **concern for abstract principles** such as human rights, social goods and wider public interest.

Figure 5.2 shows these dimensions as the Y and X axis of a grid, and shows how the seven mindsets can be mapped against these two broad ways of thinking about the world.

Figure 5.1 – The seven key mindsets



Individuals may hold two mindsets at once and express contradictory views towards the same scenario of data-sharing activity. Commercial access to health data is a very new topic for most of the general public, hence opinions are not fully formed and there is much confusion in public reactions to this new context.

5.2 Seven mindsets among public and patients

5.2.1 Community Optimists

Open to commercial interest in other aspects of life, Community Optimists accept it as necessary for social development. They are concerned to preserve citizens' rights and create social goods, so they favour non-commercial organisations to deliver public benefits in principle and are keen to avoid whittling away the remit of the NHS, for example. They are very concerned to protect vulnerable groups, but they are optimistic so are not too worried; they feel society is more likely to become more equitable and fair than the reverse.

They are optimistic about public-private partnerships and see commercial access to health data as an acceptable compromise in return for benefits and developments in medicine. Because they think about society more than about themselves, and also count the benefits they receive, they feel the trade-off between personal **loss of privacy/risk of personal harm** and **general public gain** is worthwhile.

"If it was to develop a drug that could save lives and benefit people I'd be for it."
General public, Sheffield

They do not accept commercial access to health data in which only a private company is seen to benefit, such as calculating insurance premiums.

They are the most open and optimistic about the potential for private sector involvement in the uncertain world of genetic research.

*"As a society to go forward, you have to take on board certain changes **(…)** the impact it will have at a later time, **we'll probably be okay with it…"***
General public, London

Some hold the view that citizens have a **duty to share health data** if they want to live in a society where they access public health services and receive benefits of health research. They are interested in brokering a new social contract, able to conceive of how sharing health data is a new type of transaction, where individuals and society have to accept some element of risk-taking in the name of public gain.

"If you want to benefit from the society we live in, the price is your personal information."
General public, London

This mindset group tends to include **younger participants**. Having grown up with technology, they are used to sharing their personal data online. Younger participants also have a stronger sense of abstract principles, and weigh up public and private risk against public gain. Impressionistically, this group also seem to be among the higher social grades and slightly better educated. This chimes with the quantitative finding that those with higher qualifications have more knowledge about data sharing and are more likely to agree that it is acceptable for commercial organisations to access health data for research.

This optimism is not without conditions and Community Optimists want a guarantee that wider public interest will be a key driver and outcome of any activity involving a commercial company.

"As long as it's not used to detrimental effect, it's how you find breakthroughs."
General public, Sheffield

In the quantitative survey, internet access also appears to be linked to support for commercial access, with over half of daily internet users (56 per cent) and less frequent internet users (52 per cent) supporting commercial access to health data for research. This compares with under two-fifths (39 per cent) of those who do not have internet. This could also be a function of the above differences in social grade and qualification level.

5.2.2 Community Worriers

Community Worriers see clear public benefits to sharing health data but are more wary of allowing commercial companies to get involved, believing that they might be tempted to put their own interests first. They tend to be **older and from higher social grades**, suspicious of profit-making motives and less trusting of commercial organisations overall.

They worry about the risks of personal/public harm and want to know that the companies' involvement will be overseen by a public sector body. They are also wary of privatising useful knowledge and compromising public benefit as a result.

"You can do loads of really good things but if you're selling it to a private sector, it becomes private knowledge for someone who has been paying so it's not public data that is being used for good purposes."
General public, London

They fear that outsourcing data analysis to private companies will lead to arm's length regulation, jeopardising the quality of the work and weakening any guarantee that public interest will be protected.

"I don't see how they (NHS) can really do anything if it's not in their hands? How can they stop things going wrong, how do they know what kind of service they're getting?"
General public, Swansea

Community Worriers show concern for vulnerable groups in society and patients likely to have particularly sensitive data and are very keen to avoid any risk of discrimination and personal harm to individuals.

They need to know that data-sharing activity is totally necessary and that data is not being collected gratuitously.

"There's certain info you need out there like your blood group to keep blood stocks up. So that's fine, but there's certain personal info they don't need."

General public, Sheffield

5.2.3 Pragmatic Sceptics

Pragmatic Sceptics are highly sceptical of commercial organisations' motives and the coexistence of public benefit and profit.

"The scale of this project doesn't automatically l*ead me to believe anyone on this table will benefit. They will have to recoup the money from somewhere and it* ***won't be making my life better."***
General public, Swansea

"Nowadays nothing is sacred if it's got a dollar at the end."
General public, Swansea

Their unease is compounded by a sense of inevitability and the belief it is too late to prevent or control any negative impact. They think data misuse is already a problem, lack faith in existing data protection systems and think the problem will only get worse.

"You know that people can access data that shouldn't."
General public, Dundee

Thus, though very wary of commercial involvement in health research, they sadly accept the necessary evil; they recognise data-sharing as important for health improvement but see commercial involvement as an imperfect solution.

"It's not perfect but if we want to progress things like this have to be done."
General public, Sheffield

They are often older individuals. Compared to the Community Worriers or the Community Optimists, they are much more focused on pragmatic concerns about the direct impact to them personally and the potential for personal gain or loss than on wider social and ethical considerations. Many have been hounded by direct marketing and insurance companies contacting them with deals *"for the over 60s"* and this biases their view. They are also strongly influenced by media stories of data-leaks, hacking and identity fraud.

They like the idea of giving their **permission** for a commercial company to access data as they feel that they are somehow losing their privacy and talk about lots of health data being 'theirs'. When prompted, they have unrealistic requests for opt-in consent models and respond very well to case studies where individuals are explicitly asked for their consent. This alone can drive their acceptance and they warmed to the examples of **Crowdsourcing health data from patient forums** and **Pharmacists accessing summary care records** (see chapter 4 for details of these examples).

"First of all they agreed and registered. So long as the patients know that they're sharing their information, then it's up to the individual."
Patients (non-severe conditions), Swansea

They were mainly concerned with not being involved *themselves.* So long as they had the guarantee of being able to opt-out then they were more accepting of commercial access because they felt it did not affect them.

5.2.4 Pessimistic Dystopians

Pessimistic Dystopians, on the other hand, are abstract thinkers, considering the impacts on society. In this, they are like Community Optimists. However, they are the reverse of optimistic. They are the most wary of any commercial involvement. They share similar concerns to Pragmatic Sceptics in relation to security of data and individual rights to privacy, but are not only concerned about the potential negative effect on themselves, but fear large-scale negative impact for all society. They have the strongest objections to commercial access and compared to other groups tend not to shift in their views. Being allowed to opt out as an individual does nothing to allay their fears of wider public risk.

They have very limited trust in commercial organisations and not only fear that there may not be a public benefit, but that giving companies access to health data will create new public harms.

"Are they making money by creating illnesses where there aren't any?"
General public, Swansea

They are prone to talk about the worst case scenario. They have strong values when it comes to overall social goods and see commercial involvement undermining the very principle of prioritising human rights and social care (i.e. the NHS).

"This goes against our society and values where we all pay for everyone's health. I'm proud of the UK for helping those who can't help themselves, this is being eroded."
General public, Sheffield

They sometimes overstate the negative implications for society, envisaging a nightmare 'Big Brother' reality with excessive amounts of health data sharing and commercial companies being given access rights under any circumstance.

Prone to think of the worst case scenario, they have strong red lines of unacceptability, particularly when it comes to **uncertain future uses** of data where the public benefit is not clear. For example, in the case of calculating insurance premiums they could not imagine the public benefits and pictured huge disadvantages to society.

Pessimistic Dystopians are often older and from higher social grades.

5.2.5 Fine By Me

The Fine By Me mindset represents a stark contrast to the scepticism, nervousness and fatalism of others. Overall, 'Fine By Mes' are confident about commercial influences in life, do not worry about it, and are basically pragmatic and focused on themselves rather than wider social questions. With this more relaxed approach, they tend to express a more neutral attitude towards the issue and ask fewer questions about the specifics of the activity. Many are younger and adopt a very blasé attitude, having grown up in a world characterised by commerce and data-sharing on and offline. They are **used to seeing themselves as consumers** not citizens and are *'commercial access natives'*, thus they are more at home in this new context and have fewer red lines of unacceptability.

They are more concerned with personal risk than social abstracts, but see these as limited and are not worried by any of the current examples. Many with this mindset are from lower social grades and do not consider the effect that commercial access might have for wider society; it may, for example, be this group that contribute to the unusual quantitative finding that those in the DE social grades are more likely to support health data sharing for marketing purposes, despite this group's scepticism on other measures of acceptability of commercial access (see chapter 6 for details).

With experience of data-sharing in other areas of life - or no experience and little interest - they assume that personal data will be accessed and stored safely and securely. This explains their comparatively high levels of comfort with commercial access.

"It wouldn't be just kind of, an average lady at the desk shouting out your name and address – it would be at the back. Probably."
General public, Wrexham

Even where commercial access may be used for marketing purposes, participants with this mindset do not reject it outright as they focus on nuisance marketing and junk mail not the whole concept of marketing as a "dirty" activity.

Fine By Mes are the least insistent on the requirement for public benefit, though they do see this as an ideal additional outcome of the research.

5.2.6 Patients and cohort members can have unique, but varied perspectives

The patients with severe, non-severe and rare long-term conditions in the deliberative research had often witnessed the benefits *and* risks of sharing health data, through having greater contact with health services.

Some focused on the potential benefits that data-sharing can bring and thus accept a higher level of personal risk; or expressed a different attitude to risk having faced life-threatening situations. They recognise that they can receive better direct care as a result of personal information about them being shared between public and private bodies and there is huge potential for others with their condition to benefit.

"Doesn't really bother me where they put it, I had a brain haemorrhage at 21, if this can help others like me they can do whatever."
Patients (rare conditions), Sheffield

Cohort members had the additional feature of having shared health data from a very young age. Like the Fine By Mes, they were more relaxed about the subject and, having had positive experiences, they often adopted an optimistic mindset. Many knew that health data-sharing could be carried out in a supportive environment and explained how helpful they found the constant reassurances and provision of information that researchers conducting the study gave them. This highlights the significance that informing people and explaining the "WHY" behind data sharing and research activity has in shaping public views.

Patients stand to gain more from data-sharing but also have more to lose through identification (e.g. discrimination by employers, stigma among colleagues) and contrasting attitudes are noticeable. Some patients warned of the negative repercussions they may face if their personal health details end up in the wrong hands. Some patients told stories of where they had been treated unfairly, in many cases by public bodies or employers, as a result of information being passed on. These patients tend to have very stringent expectations of health data security processes and want to be able to trust their GP and other healthcare professionals to keep data private. People with rare diseases were particularly vocal on the potential for negative impacts, financial, social and emotional; for example being discriminated against at a job interview, refused a bank loan or having trouble finding a life partner.

"It's more likely with rare conditions that they can identify us, because there are only a certain number of people with that condition. If it's diabetes or heart disease, something that lots of people have then that's different, but with rare diseases you stand out more clearly."
Patients (rare conditions), Sheffield

Two distinct mindsets emerge from discussions with patients and cohort members: **'Monitor Me', and 'Fed Up'**. A slightly different tension exists for these participants, between pragmatic challenges of having their very sensitive and potentially highly identifiable health data shared outside of the public sector, and the life-saving benefits that either they or others may receive. The Monitor Me mindset also applies to cohort members, many of whom spoke about how willing and happy they were to be subjects of health research and data sharing. This was supported by their attendance and contributions at the workshop itself, where many expressed a genuine appreciation of the value that data sharing can have in the field of health.

For Monitor Mes, the life-saving effect that health data sharing can have makes commercial access more acceptable. The potential benefits to either themselves or future generations outweigh the risks that they may personally face of being identified, or discriminated against. Their experience of taking part in health research or their understanding of how necessary health research is for medical advances makes them more comfortable with the activity.

"I just recently gave my consent to research because I've had cancer, as it's all to help the future, my daughters. And my name will stay on all the blood test records so it can help future generations from my family."
Patients (rare conditions), Sheffield

These people are also more aware and pragmatic about the interplay of commercial and public sector organisations in healthcare and they therefore do not have the same level of surprise or discomfort at a commercial companies involvement. On the contrary, they are often able to see commercial involvement as a positive means to improving services and finding cures for diseases, given their financial resource. Moreover, many can recall tangible benefits for themselves, as the involvement might lead to improvements in their direct care.

"Speaking from personal experience I'm on dialysis. They've got lots of information on renal research. It's great because I know if I go in for tests, I can check the next day on the website... that's fantastic – statistics for me are a cure."
Patients (non-severe conditions), Glasgow

Quite unique to this group, is a sense of **personal responsibility**, the idea that they have a duty to allow their health data to be used for research. They can feel the urgency of health data sharing and the crucial hope that it gives for medical advancement.

"I'd look at that and say well it might not help me, I'm gone too far, but it might help the person that's coming up behind me with the same illness, stop them going through what I have."
Patients (rare conditions), Sheffield

Monitor Mes can extend their optimism to other members of the general public and project their view that so long as individuals and families benefit then all types of health research should take place regardless of whether they involve a commercial organisation. This even applies to genetic research where future uses and implications are unknown but the potential for life-saving impact is high.

"I don't think people would have a problem with (sharing) even DNA, if it was used for the betterment of humanity – if it's to save a child's life, make someone's life more comfortable, prevent pain... that's wonderful."
Patients (rare conditions), Sheffield

The flip side of greater contact with the health service is that some patients can feel over-monitored and frustrated at what some see as constant demands for them to share very personal and sensitive information. This group we have called the **Fed Ups.** This feeling usually relates to their experience of direct care and the sometimes repetitive nature of updating different healthcare professionals with information about their health, but it can radically colour their views of health data-sharing more generally.

"My son has a medical file inches thick - but every time we go to a doctor we have to spend 30 minutes briefing them about what's wrong, what they should already know, before we even get to the thing we're there for and want to talk about – and you can see them rolling their eyes, you can tell they haven't seen the notes before."
Patients (rare conditions), Sheffield

Fed Ups express feelings of upset and even anger which are compounded when they consider the involvement of a commercial company who they see as yet another party trying to access information about them.

Some express specific unease because they feel they are being classified by the information they give. They question the accuracy of these classifications and the implications they can have for their relationships in other spheres – personal and professional. This can cause them to call into question the whole value of collecting and sharing patient data.

***"I've been terribly treated by people over the years, who have led** with the mental health problem instead of the physical."*

Patients (rare conditions), Sheffield

Many make the further point that health data-sharing as it stands is not a two-way process – while it often benefits commercial organisations, they do not see the benefits to the NHS or themselves.

*"I went to a meeting about diabetic care, and they all just talked about data, never mentioned the words patient care – **that smacks of control; 'oh we've got the data, we know everything's gonna be alright'."***

Patients (severe conditions), London

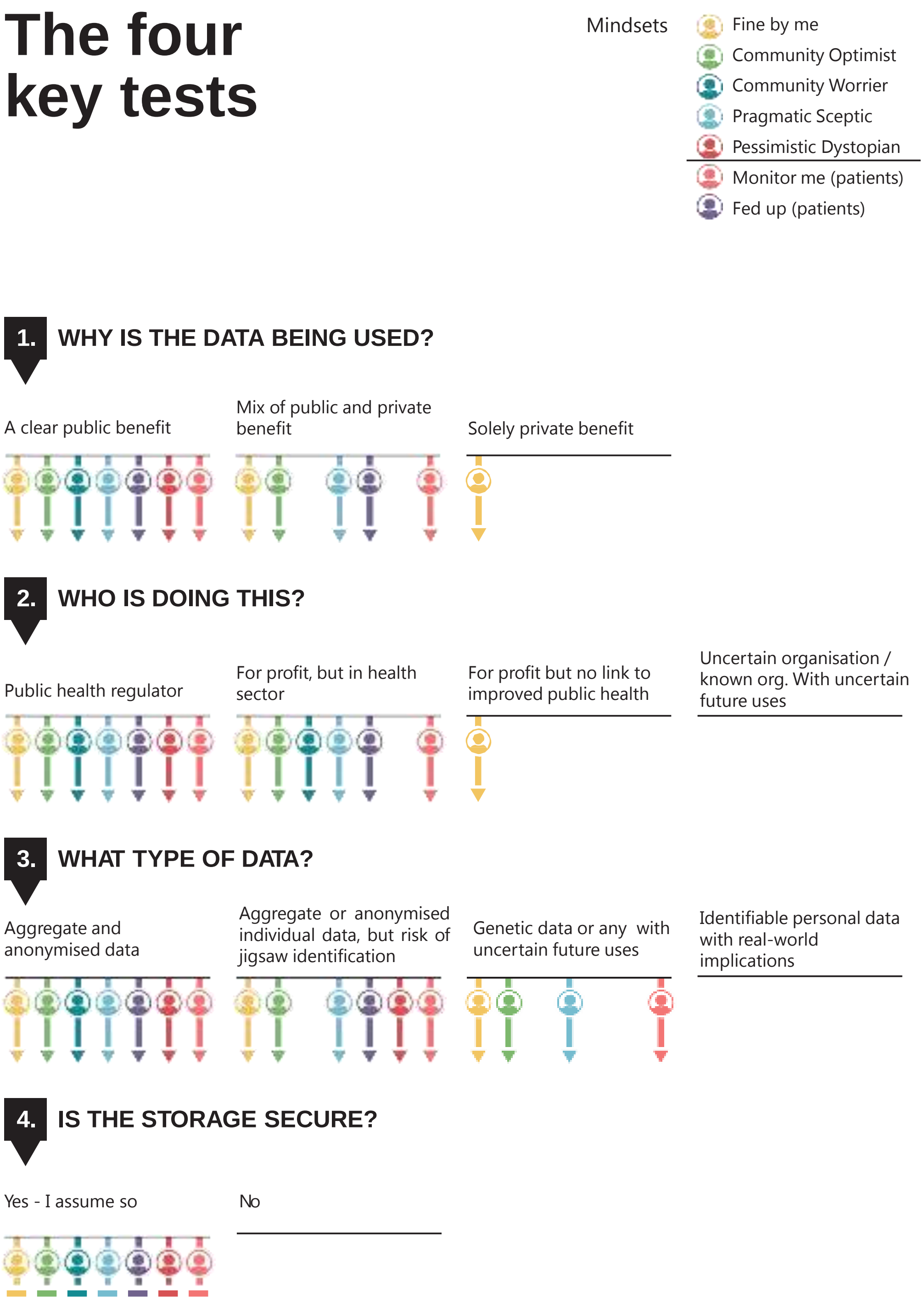
Rare disease patients were often the most likely to display the Fed Up mindset. Several of them had developed a very suspicious attitude towards commercial organisations accessing their data, after having (or hearing about) bad experiences of commercial companies exploiting their rare health status and marketing products to them.

These concerns reflected views of many **healthcare professionals** in the research, who stressed that there is a pressing need to protect those individuals with most to lose from commercial access (i.e. rare disease sufferers).

5.3 Influence of mindsets on approaches to the four 'key tests' and red lines of commercial access

The mindsets participants identified with had a strong influence upon their approach to weighing up the potential value and risks associated with the different examples of commercial access to data, and to the way in which they applied the four 'key tests' to each example. Linked to this, the different mindsets caused people to draw 'red lines' under different circumstances and at different levels of commercial access. For example, while all mindsets were generally accepting of commercial access with a clear public benefit, for Pessimistic Dystopians and Community Worriers the test of 'who benefits' is failed as soon as there are benefits to the private sector. Taking this a step further, only those maintaining the 'Fine by me' mindset are able to accept commercial access which solely benefits private sector organisations. The interaction between the mindsets and the four key tests is illustrated in figure 5.2.

Figure 5.2 – The seven mindsets and the four 'key tests'



6 Findings from the quantitative survey

6 Findings from the quantitative survey

Following the deliberative workshops, Ipsos MORI conducted a face-to-face survey of adults in Great Britain to collect quantitative findings about their attitudes towards commercial access to health data. This chapter is a summary of the findings from that survey. The full topline findings can be found in the appendix.

The questionnaire was designed in conjunction with the Wellcome Trust and the advisory group to probe on some of the concepts and ideas that came out of the qualitative stage. The survey was designed to build on previous quantitative surveys in this area that have broadly found there to be a tentative support for using data for research[14], but also genuine concerns about key factors, such as:

- The public's particular perception of privacy associated with health and medical records[15];
- The context of the particular health data sharing playing a key role in determining trust and acceptability[16]; and
- The specific problem of sharing health data between the public and private sectors[17].

The design of the questionnaire took into account interesting parts of the qualitative findings that appeared to warrant and suit quantitative follow-up. However, at the same time, it was appreciated that priorities needed to be made about could be asked in the allotted time, and that the exercise of conducting quantitative research is very different from a qualitative project, where participants are introduced in depth to the context across an entire day. It was also clear that the complexity of parts of the qualitative research could not be replicated in the survey due to the constraints of space and methodology.

Ipsos MORI interviewed a representative quota sample of 2,017 adults across Great Britain aged 16 and over. Face-to-face interviews were conducted in-home between 30 November and 11 December 2015. Data are weighted to the known population profile.

---

[14] Ipsos MORI, Royal Statistical Society (2014), 'New research finds data trust deficit with lessons for policymakers', https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-lessons-for-policymakers.aspx (accessed 06/12/15)

[15] Ipsos MORI, Joseph Rowntree Reform Trust (2014), 'Poll on privacy and data sharing for the Joseph Rowntree Reform Trust', https://www.ipsos-mori.com/researchpublications/researcharchive/3388/Poll-on-privacy-and-data-sharing-for-The-Joseph-Rowntree-Reform-Trust.aspx (accessed 18/01/16)

[16] Ipsos MORI, BIS (2014), 'Public Attitudes to Science 2014', https://www.ipsos-mori.com/researchpublications/researcharchive/3357/Public-Attitudes-to-Science-2014.aspx (accessed 18/01/16)

[17] Ipsos MORI, Royal Statistical Society (2014), 'New research finds data trust deficit with lessons for policymakers', https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-lessons-for-policymakers.aspx (accessed 06/12/15)

6.1 Key findings

- **More support than oppose health data being used by commercial organisations undertaking health research** - over half (54 per cent) support commercial access to health data for research, while a quarter (26 per cent) oppose it. This leaves a fifth of respondents who either said that they neither support nor oppose commercial access to health data (19 per cent) or saying 'Don't know' (2 per cent).

- **Awareness of health data usage is low -** just one third (33 per cent) have heard a great deal or a fair amount about how the NHS is using health data. This detailed awareness falls to 16 per cent for commercial organisations and 18 per cent for academic researchers.

- **Faced with losing out on research, people will opt for this research being done by commercial organisations** - a majority of people (61 per cent) would rather see commercial access to health data happen than lose out on the benefits that research involving these organisations can bring. A quarter (25 per cent) would still rather that research did not happen if commercial organisations had to have access to the data.

- Permission is still a key part of people's attitudes towards their health informa**tion** – the desire for the NHS to ask permission before this kind of data sharing takes place is widespread, with 54 per cent saying that they would rather the NHS ask patients' permission before sharing data with commercial organisations, even if this meant that without permission the research cannot take place. The survey also found that a majority (53 per cent) want to see strict rules in place that data could not be passed to third parties.

- **Sharing health data for the purposes of insurance or marketing both face significant resistance from the public** - just a quarter (26 per cent) support sharing anonymised health records with insurance companies so they can develop their insurance prices. Support for companies marketing health products using anonymised health records is higher, with 37 per cent supporting this purpose, but still lower than some of the support for commercial organisations generally using health data.

- **There is still a core segment who cannot see any circumstances under which they would allow commercial organisations access to NHS health data -** 17 per cent of people say they would not want commercial organisations to have access to health data for research under any circumstances. Of these, one-fifth (20 per cent) say commercial organisations cannot be trusted to store the data safely, and a similar proportion say that profit should not be made from health data, even if there are potential societal and health benefits as well.

- A "social" value to data resonates, but many still see a "financial" value to health **data -** the majority agree that health data has financial value (50 per cent) and societal value (67 per cent).

- **Having a safeguard in place makes a difference, whatever the safeguard is** - using a split sample experiment, the survey found that there is no silver bullet safeguard that can restore public trust (between 56 per cent and 64 per cent agree with data sharing with any of the named safeguards in place). However, knowing that a safeguard is in place, regardless of the type of safeguard, makes a difference to acceptability of health data sharing with commercial organisations (agreement is 49 per cent for those respondents where no safeguard was mentioned).

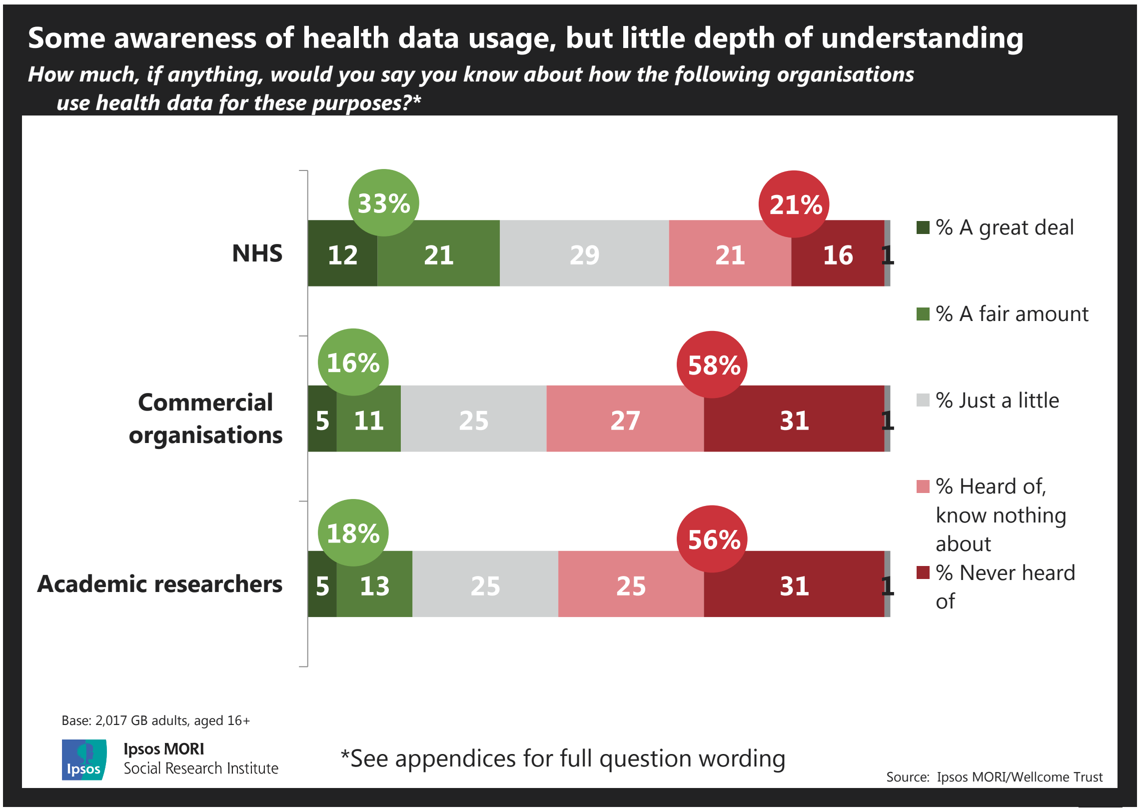
- **Despite this, people are ready to insist on many different conditions to safeguard their data -** of various different conditions that could be placed on commercial organisations to allay fears for the public, strict rules about not passing data on (53 per cent); all names and personal information being removed (52 per cent); and storage of the data in a secure facility (47 per cent) were the most popular. However, there is no clear preference for respondents on which conditions should be in place, and over two-fifths (43 per cent) want 3 or more of these conditions in place.

6.2 Awareness of health data usage

The survey began by assessing awareness of the usage of health data within different organisations: the NHS, commercial organisations and academic researchers. Health data was defined as data 'collected from patients in hospitals and GP practices', which has had personal information, such as name and address, removed.

Many have heard of health data sharing by these organisations. Over four-fifths (83 per cent) say they have at least heard of health data sharing in the NHS. The proportion saying they've at least heard of data sharing with commercial organisations and academic researchers is lower (68 per cent each). Despite this surface-level understanding, more detailed awareness of data usage by all of these organisations is low.

Figure 6.1 – Awareness of health data usage



Some awareness of health data usage, but little depth of understanding

How much, if anything, would you say you know about how the following organisations use health data for these purposes?*

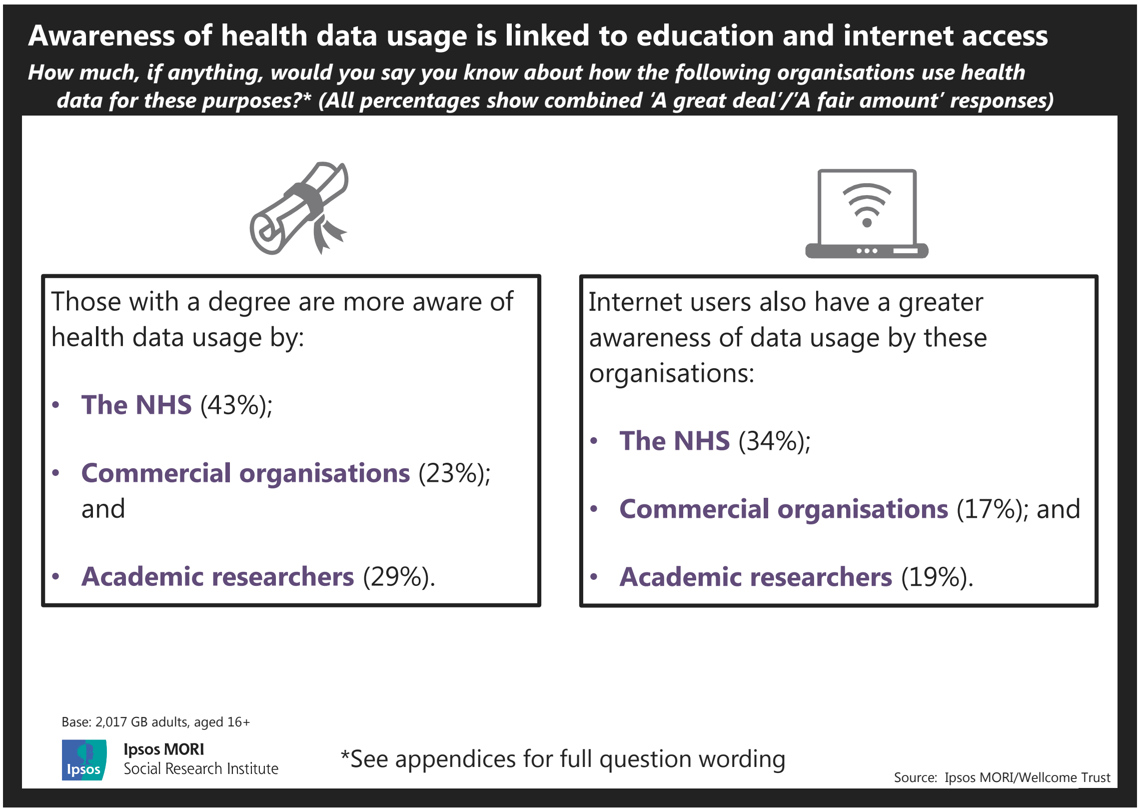
Base: 2,017 GB adults, aged 16+

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

Educational attainment and internet use appear to be linked to awareness of data usage, as outlined in figure 6.2.

Figure 6.2 – The link between education, internet access and awareness of health data usage



Awareness of health data usage is linked to education and internet access

How much, if anything, would you say you know about how the following organisations use health data for these purposes?* (All percentages show combined 'A great deal'/'A fair amount' responses)

Those with a degree are more aware of health data usage by:

- **The NHS** (43%);
- **Commercial organisations** (23%); and
- **Academic researchers** (29%).

Internet users also have a greater awareness of data usage by these organisations:

- **The NHS** (34%);
- **Commercial organisations** (17%); and
- **Academic researchers** (19%).

Base: 2,017 GB adults, aged 16+

Ipsos MORI
Social Research Institute

*See appendices for full question wording

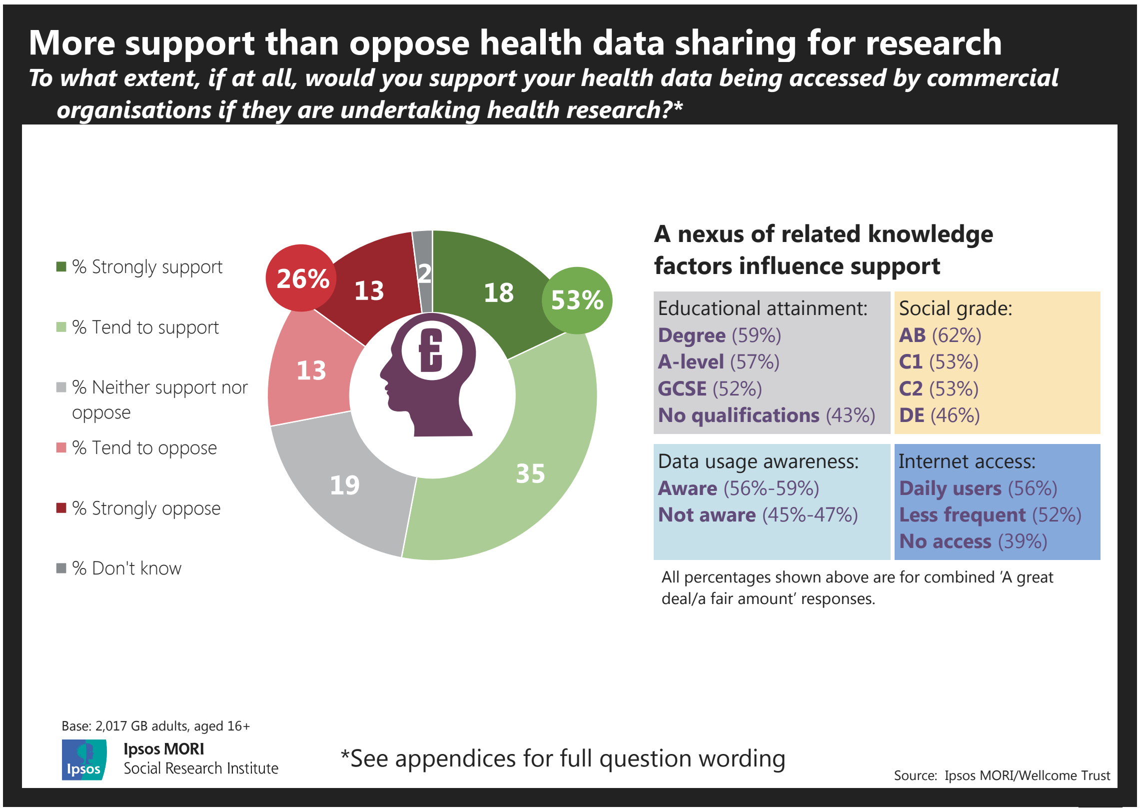
Source: Ipsos MORI/Wellcome Trust

6.3 Support for commercial access to health data

Respondents were then asked to what extent they support commercial organisations having access to their health data for the purposes of undertaking health research. It was explained that health data is information that 'is gathered about individuals by the NHS when they receive care' and that this information would not include name, address or contact details.

More support than oppose sharing health data with commercial organisations for the specific purposes of health research. Over half (54 per cent) support commercial access to health data for research, while a quarter (26 per cent) oppose it. This leaves a fifth of respondents who either say that they neither support nor oppose commercial access to health data (19 per cent) or saying 'Don't know' (2 per cent). Some of the key factors related to support are displayed in figure 6.3.

Figure 6.3 – Support and opposition for health data sharing for research purposes



More support than oppose health data sharing for research

*To what extent, if at all, would you support your health data being accessed by commercial organisations if they are undertaking health research?**

A nexus of related knowledge factors influence support

Educational attainment:
Degree (59%)
A-level (57%)
GCSE (52%)
No qualifications (43%)

Social grade:
AB (62%)
C1 (53%)
C2 (53%)
DE (46%)

Data usage awareness:
Aware (56%-59%)
Not aware (45%-47%)

Internet access:
Daily users (56%)
Less frequent (52%)
No access (39%)

All percentages shown above are for combined 'A great deal/a fair amount' responses.

Base: 2,017 GB adults, aged 16+

Ipsos MORI
Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

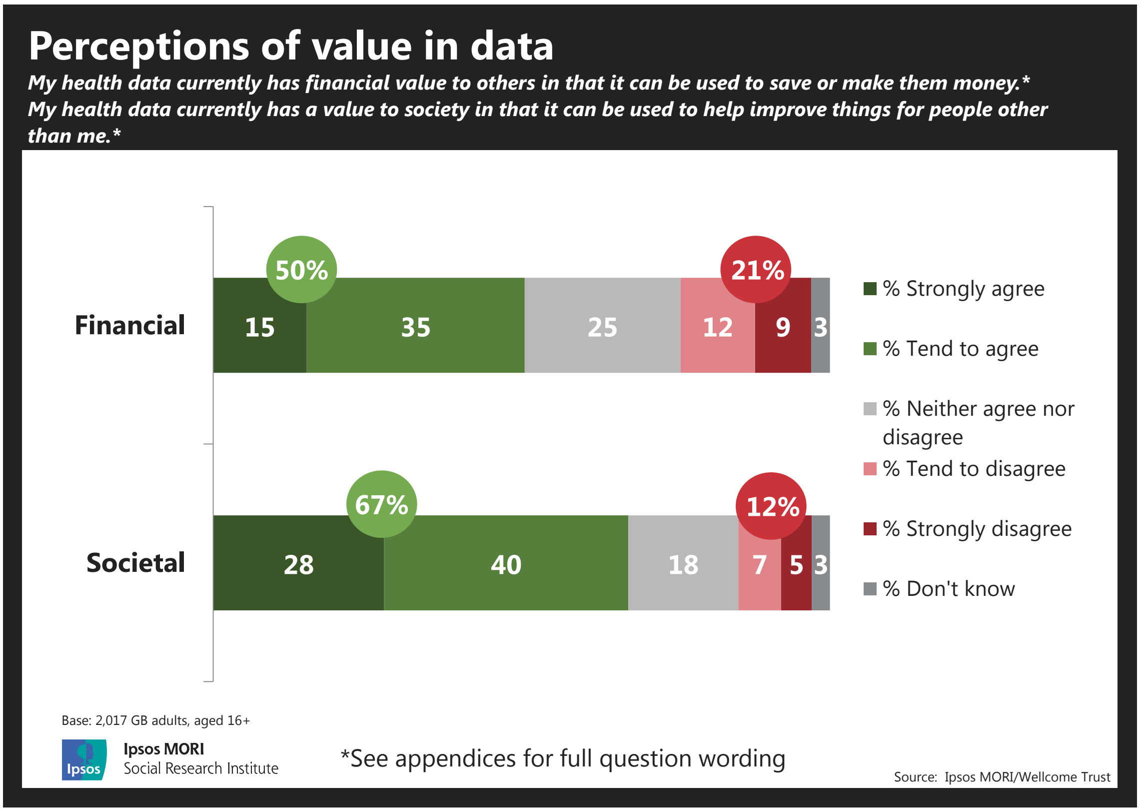
Those who later in the questionnaire reported an agreement that there is value in health data, of either a financial or societal kind, are more likely to support commercial access to health data than those who do not believe this value exists. Two thirds (64 per cent) of those who think there is a financial value to data also support commercial access to health data, compared with 43 per cent of those who not believe there is a financial value to health data. Similarly, seven-tenths (68 per cent) of those who agree that data holds a societal value, also support commercial access to health data, compared with a quarter of those who disagree that there is societal value in health data (26 per cent).

6.4 Value in health data

Respondents were asked to what extent they agree or disagree that two different types of value of health data exist – financial and societal value. Financial value was described as 'financial value to others in that it can be used to save or make them money', while societal value was described as being 'used to help improve things for people other than me'.

Half of respondents agreed that health data has financial value to others (50 per cent), while just a fifth (21 per cent) disagree with this statement. There is more widespread agreement that data holds a value to society, however, with over two thirds agreeing (67 per cent), compared with 12 per cent who disagree. Similarly to some of the other questions, many are not willing to express an opinion either way when it comes to data having these kinds of value. A quarter (25 per cent) neither agree nor disagree that health data has financial value to others, while nearly a fifth (18 per cent) responded in this way when asked whether health data has a value to society.

Figure 6.4 – Perceptions of value in data



Perceptions of value in data

*My health data currently has financial value to others in that it can be used to save or make them money.**
*My health data currently has a value to society in that it can be used to help improve things for people other than me.**

Base: 2,017 GB adults, aged 16+

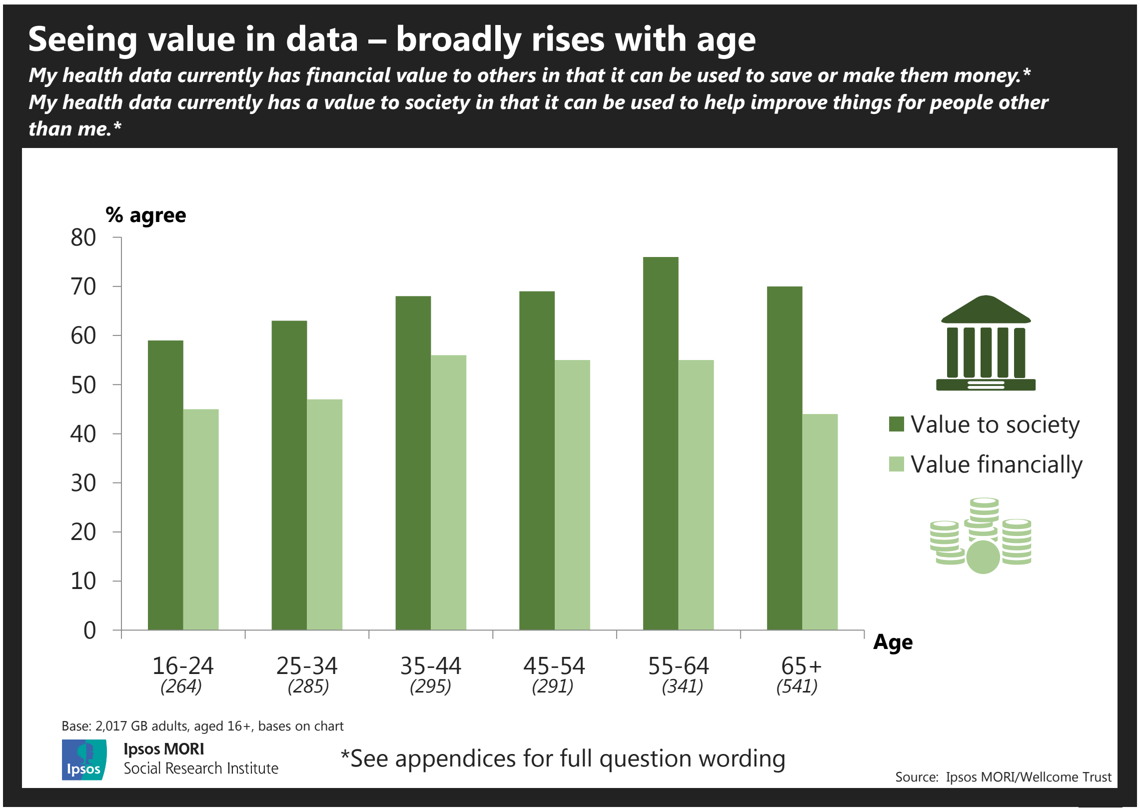
*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

There are differences in the way sub-groups attach value to health data. Men (55 per cent) are more likely than women (45 per cent) to agree that health data has a financial value, and men are also likely to believe that health data has a value to society (70 per cent of men agree, compared with 65 per cent of women).

The findings also demonstrated a significant relationship between the two types of value being perceived in health data and age, as illustrated in figure 6.5.

Figure 6.5 – Perceptions of value in data – patterns by age



Seeing value in data – broadly rises with age

My health data currently has financial value to others in that it can be used to save or make them money.*
My health data currently has a value to society in that it can be used to help improve things for people other than me.*

Base: 2,017 GB adults, aged 16+, bases on chart

Ipsos MORI
Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

It was also possible to see the relationship between these two statements to understand whether individuals who see the financial value of health data also see the societal value, or whether the opposite is true. As might be expected, the majority of those who agreed with data having a societal value, also agreed that health data holds a financial value to others (65 per cent), while just 17 per cent of those who agreed with there being societal value, *disagreed* that there was financial value. This suggests that these two types of value are linked, and if a person ascribes to the idea that health data has a societal value, they are also likely to believe that it has a financial value. However, the qualitative research found that while concepts of value both financial and societal made sense to participants, these ideas were not top of mind (see chapters 3 and 4).

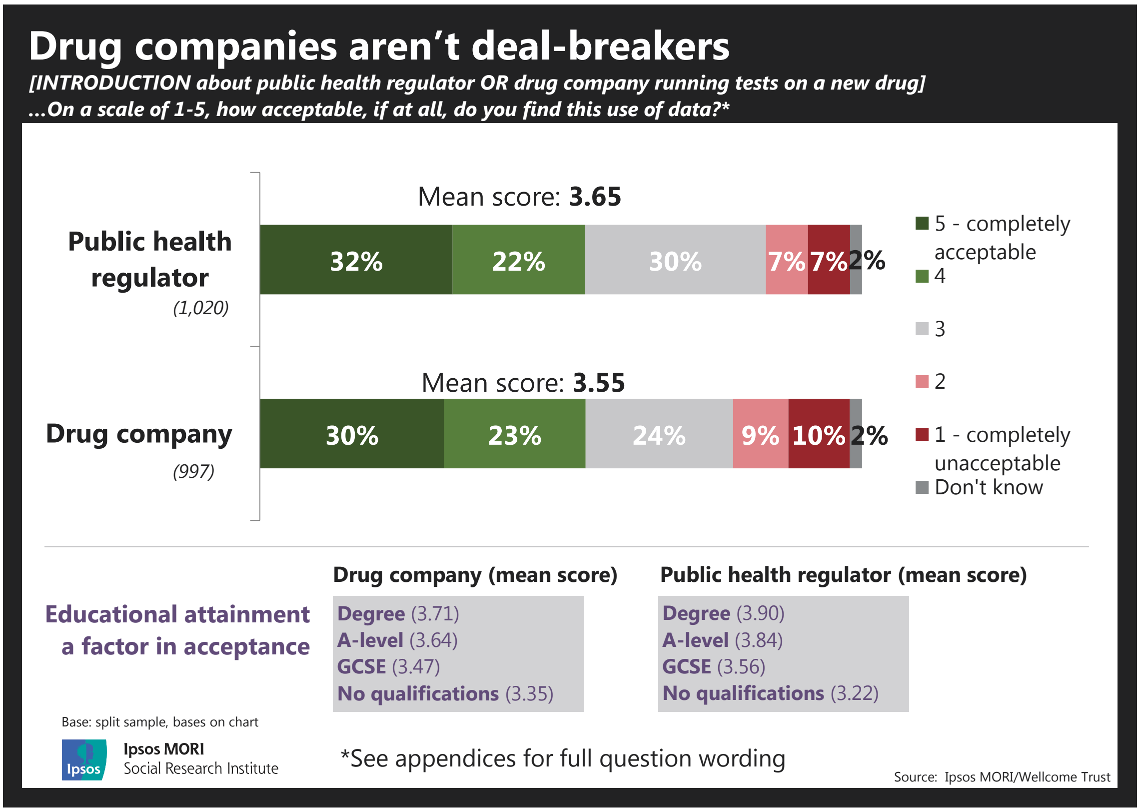
6.5 Public health regulator compared with a pharmaceutical company

Participants were shown a scenario where an organisation – either a public health regulator or a pharmaceutical/drug company – are running further tests on an approved drug to look for unwanted side effects. The participants were informed that the regulator or the company requests this information from a government department, which includes medical records with names, addresses and personal information removed. Participants were randomly allocated either the scenario with the public health regulator conducting this analysis, or the drug company (in this case, they were also informed that the company covers the costs of the analysis), and asked to rate the scenario on a scale of 1-5, where 1 was completely unacceptable and 5 was completely acceptable.

Mean scores for the two different scenarios were calculated and these showed that there is no significant difference in acceptability when a public health regulator is conducting this research as to when a pharmaceutical company is. The mean acceptability score for a drug company is 3.55 (on a scale of 1-5), while it is 3.65 for a public health regulator. There is, however, a marginally higher number who rate the drug company conducting this research as 'completely unacceptable' (10 per cent), than when the public health regulator is named (7 per cent).

Acceptability for both of these organisations running these tests was again related to educational attainment as shown below.

Figure 6.6 – Acceptability of a public health regulator vs. a drug company running research



Drug companies aren't deal-breakers

[INTRODUCTION about public health regulator OR drug company running tests on a new drug] …On a scale of 1-5, how acceptable, if at all, do you find this use of data?*

Educational attainment a factor in acceptance

	Drug company (mean score)	Public health regulator (mean score)
	Degree (3.71)	**Degree** (3.90)
	A-level (3.64)	**A-level** (3.84)
	GCSE (3.47)	**GCSE** (3.56)
	No qualifications (3.35)	**No qualifications** (3.22)

Base: split sample, bases on chart

Ipsos MORI Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

Mean acceptability for the drug company conducting this research is higher among older age groups (3.90 is the average score for 55-64s, and 3.76 for the over-65s), while younger people have lower average scores (3.36 for 16-24s; 3.25 for 25-34s; 3.49 for 35-44s; and 3.44 for 45-54s). This pattern is not repeated with the public health regulator version of the question, where there are no differences except 45-54s score a slightly higher average (3.82 acceptability) than 16-24s (3.53).

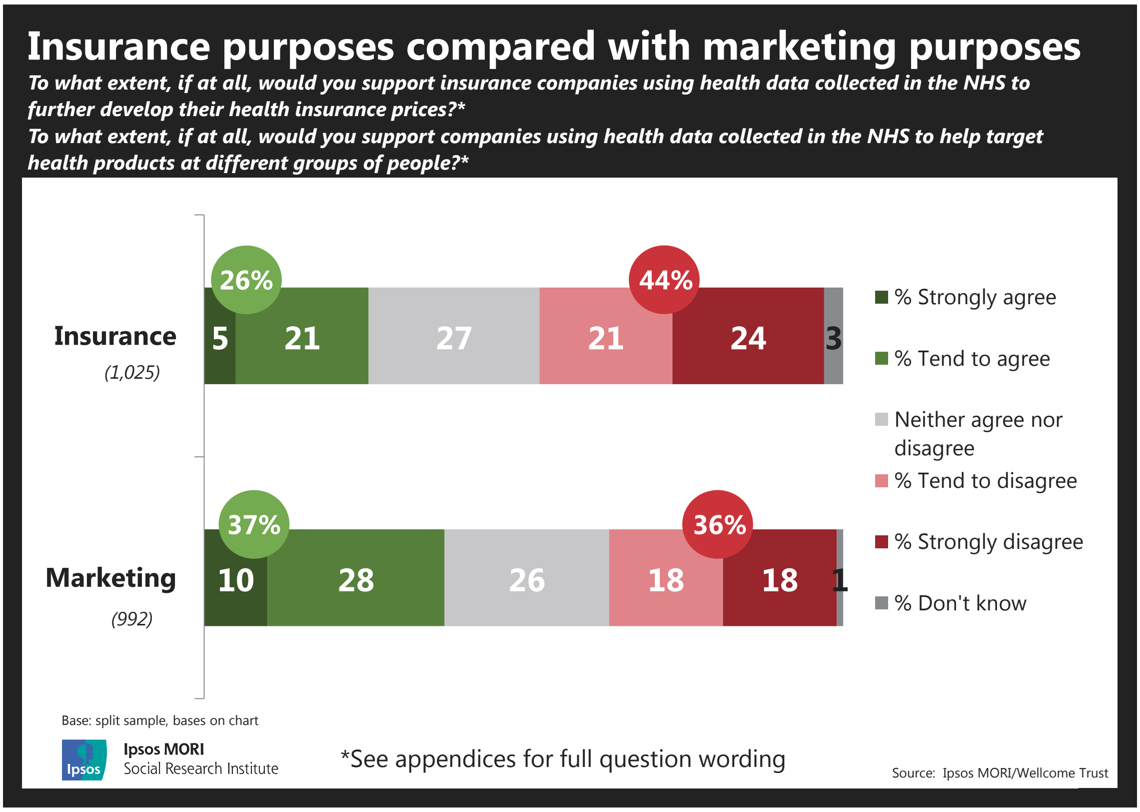
Respondents who previously supported commercial access to health data are also supportive of the drug company conducting this research (scoring an average of 3.99), while those who opposed commercial access give this scenario a lower score (2.94). Those who support commercial access to health data are equally as likely to support drug companies conducting this research as public health regulators (4.05 vs. 3.99), but those that oppose commercial access score the public health regulator (3.20) higher than the drug company (2.94).

6.6 Insurance compared with marketing purposes

Participants were randomly allocated to one of two questions about support for health data being used for different commercial purposes. Half of the sample was asked to give their support or opposition for insurance companies using health data from the NHS to develop their insurance pricing. Respondents were informed that personal information is removed from the data, and that the purpose of this process would be to tailor health insurance prices to reflect the risk of ill-health for people living in different local areas.

Opposition to health data being shared for developing insurance premiums is widespread with many more opposing (44 per cent) than supporting (26 per cent). There is also over a quarter of people (27 per cent) who sit on the fence for this scenario.

Figure 6.7 – Support and opposition for data use by insurance and marketing companies



Insurance purposes compared with marketing purposes

To what extent, if at all, would you support insurance companies using health data collected in the NHS to further develop their health insurance prices?*

To what extent, if at all, would you support companies using health data collected in the NHS to help target health products at different groups of people?*

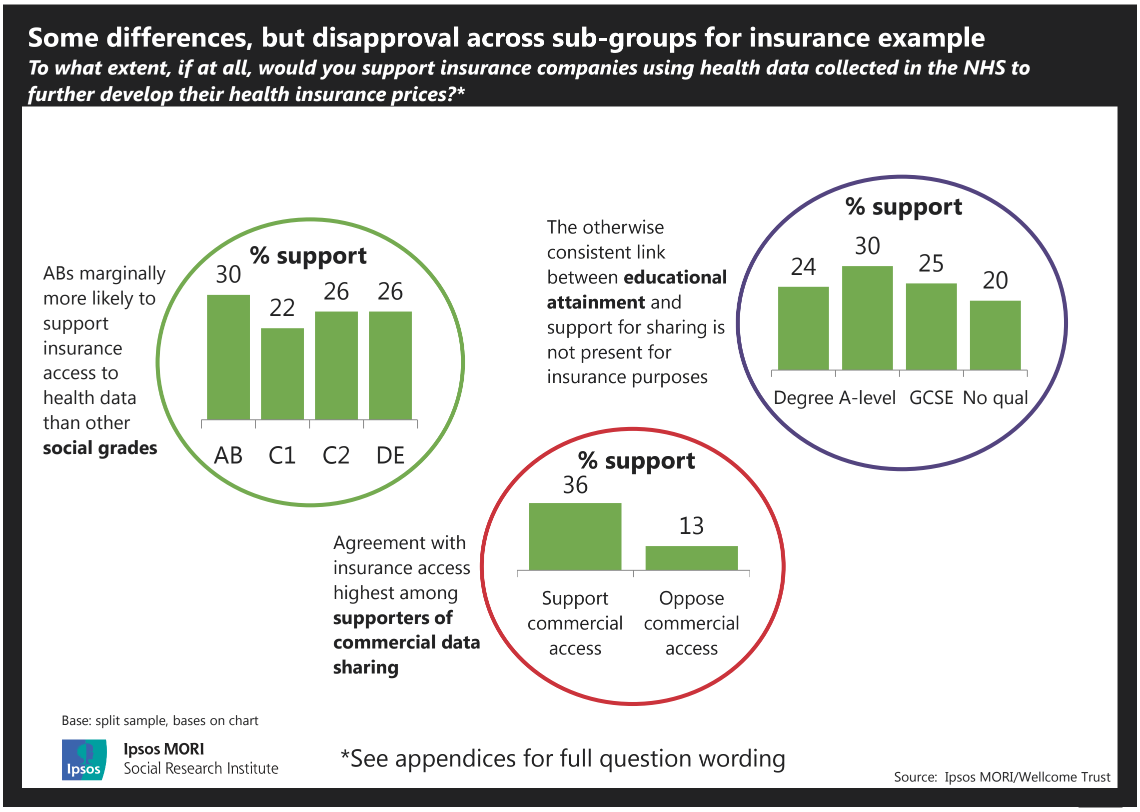
Base: split sample, bases on chart

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

Support is low across the different sub-groups for this scenario, and some of the trends in support that could be seen in other questions cannot be seen here. Those with a degree, for instance, are not significantly more or less likely to support this scenario (24 per cent) as those with no qualifications (20 per cent). Some of these differences are illustrated in figure 6.8.

Figure 6.8 – Differences in views between sub-groups – insurance companies using data



Some differences, but disapproval across sub-groups for insurance example

To what extent, if at all, would you support insurance companies using health data collected in the NHS to further develop their health insurance prices?*

ABs marginally more likely to support insurance access to health data than other **social grades**

The otherwise consistent link between **educational attainment** and support for sharing is not present for insurance purposes

Agreement with insurance access highest among **supporters of commercial data sharing**

Base: split sample, bases on chart

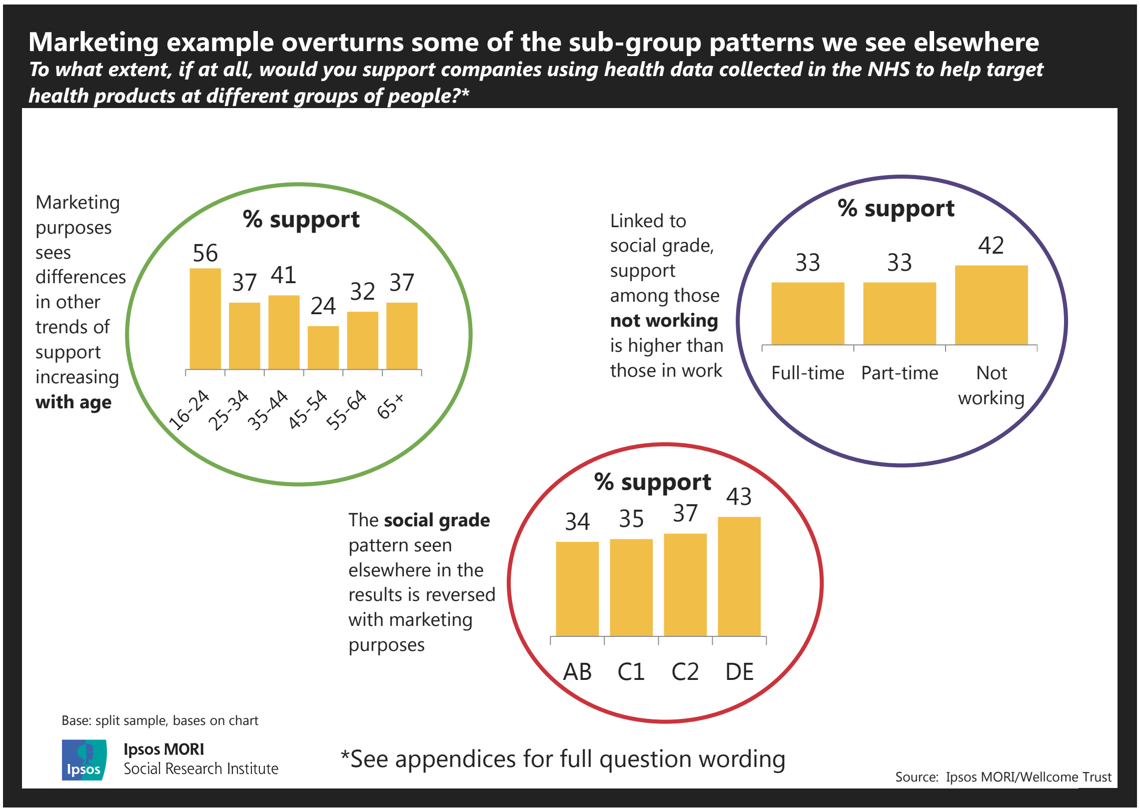
*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

The other split sample were given a similar scenario about support for health data being used for marketing purposes, such as targeting a health product – in this case a low-fat margarine – in local areas that might benefit from this. Respondents were told that the data was de-identified and were given the example of promoting a low-fat margarine in areas where there is a higher risk of heart disease.

Respondents were again told that personal information would be removed from the data. Opinion is evenly split over health data being used for marketing purposes, with a third (37 per cent) supporting and a similar proportion (36 per cent) opposing. Notably, a quarter of respondents (26 per cent) neither support nor oppose this purpose. While opposition to data being used for the above marketing example is widespread, it receives more support than health data being used for insurance pricing. Some of the sub-group differences for attitudes towards marketing purposes are displayed in figure 6.9.

Figure 6.9 – Differences in views between sub-groups – marketing companies using data



Marketing example overturns some of the sub-group patterns we see elsewhere

To what extent, if at all, would you support companies using health data collected in the NHS to help target health products at different groups of people?*

Marketing purposes sees differences in other trends of support increasing **with age**

Linked to social grade, support among those **not working** is higher than those in work

The **social grade** pattern seen elsewhere in the results is reversed with marketing purposes

Base: split sample, bases on chart

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

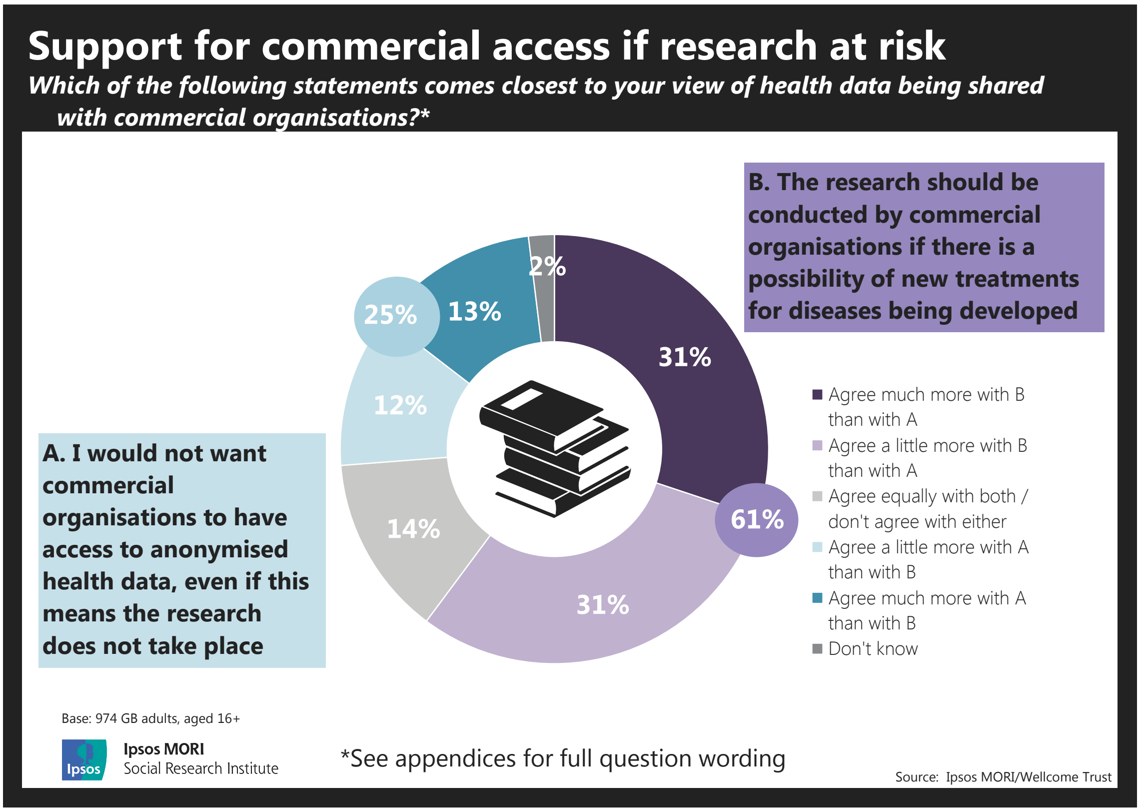
6.7 Trade-off between commercial involvement and research taking place

The sample was split once more with half of respondents allocated to a question with an introduction detailing the benefits of research for the health system, while the other half were given a question about permission in research (Q7A and Q7B – see appendix for survey topline results and full question wording). The introduction briefed respondents that it is sometimes necessary to involve commercial organisations in order to conduct research, and these organisations may make a profit from it. Respondents were then asked if they agreed more with a statement that said that commercial organisations should not have access to health data, even if this means the research might not take place, or whether the research should be conducted by commercial organisations. This question was primarily to identify whether the profit-motive overrides the benefits that research brings for participants making a judgement about how their health data is used.

Three-fifths (61 per cent) agree that research should be conducted by commercial organisations if there is a possibility of new treatments being discovered. This is more than twice the number that agreed more with the antithetical statement: that commercial organisations should not have access to health data (25 per cent). This compares with the similar, but lower, figure of 54 per cent who support commercial access to health data for research.

Over two-fifths (44 per cent) of those respondents who opposed commercial access to health data earlier in the survey, agreed here that research should be conducted by commercial organisations if, otherwise, the research would not take place. This potentially shows that when faced with the stark possibility of health research (particularly in developing new treatments) not taking place, people are more willing to allow commercial organisations to undertake this research.

Figure 6.10 – Support for commercial access if research would not go ahead without it



Support for commercial access if research at risk

Which of the following statements comes closest to your view of health data being shared with commercial organisations?*

B. The research should be conducted by commercial organisations if there is a possibility of new treatments for diseases being developed

A. I would not want commercial organisations to have access to anonymised health data, even if this means the research does not take place

Base: 974 GB adults, aged 16+

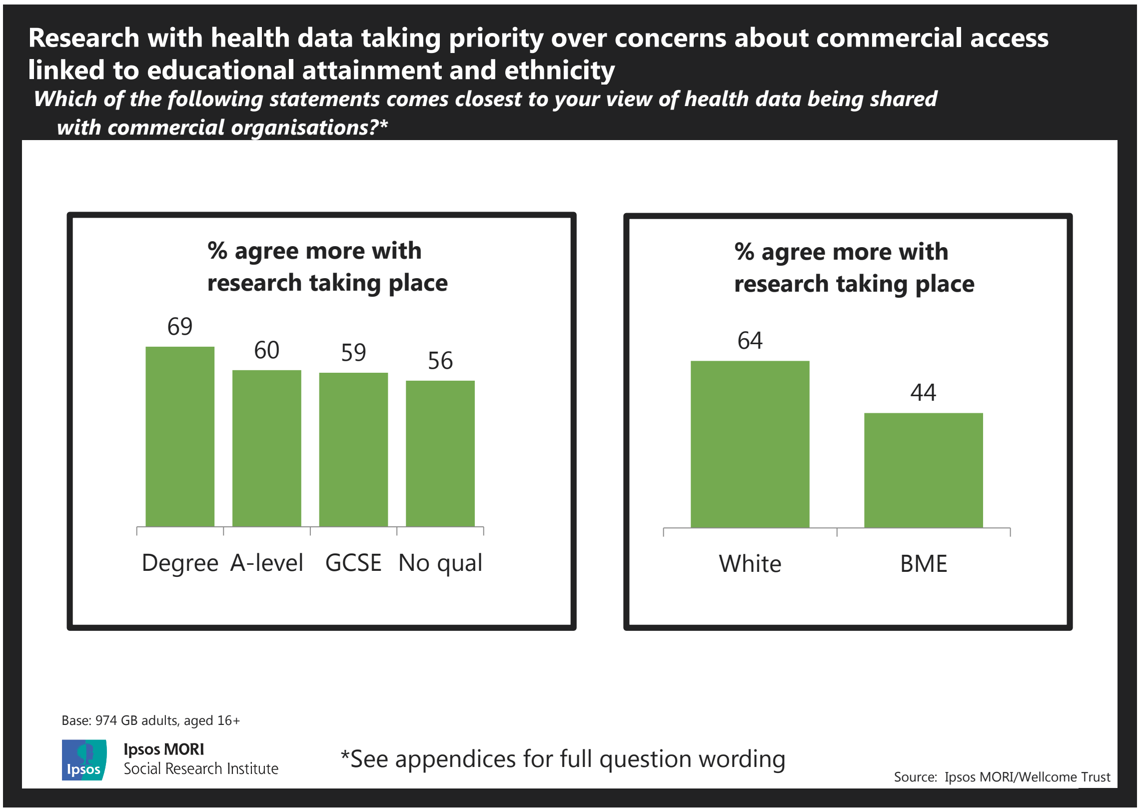
Ipsos MORI Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

As elsewhere in this research, age has a peculiar relationship with supporting commercial access. In particular, seven-tenths of 55-64 year-olds (69 per cent) agree with commercial organisations conducting research using health data, compared with a lower proportion of 16-24s and 35-44s (55 per cent and 56 per cent, respectively). Again, however, it is possible to see here that those of a younger age are also more likely to ascribe to the middle position (17 per cent of 16-24s and 25-34s; 20 per cent of 35-44s, compared with 11 per cent of 45-54s; 8 per cent of 55-64s and 12 per cent of over-65s). Other factors that appear to influence coming down on the side of research being conducted by commercial organisations are shown in figure 6.11.

Figure 6.11 – Support for commercial access if research would not go ahead without it; sub-group differences



Research with health data taking priority over concerns about commercial access linked to educational attainment and ethnicity

Which of the following statements comes closest to your view of health data being shared with commercial organisations?*

Base: 974 GB adults, aged 16+

Ipsos MORI Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

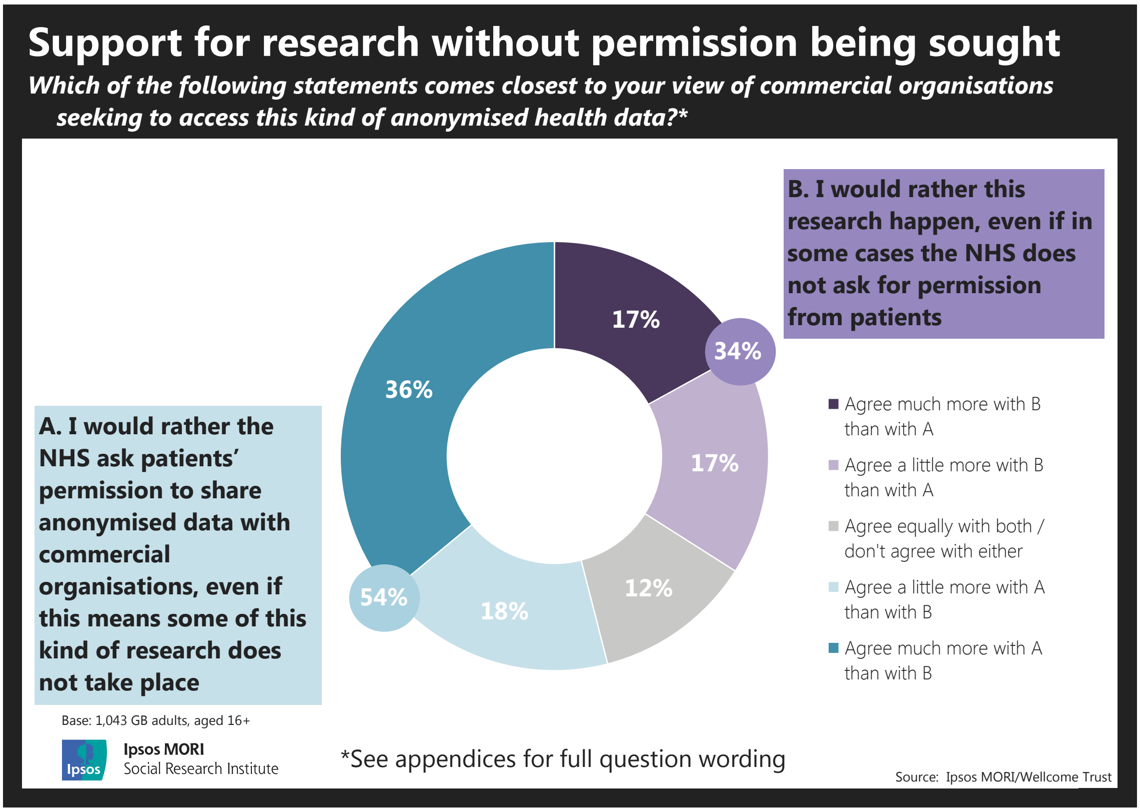
6.8 Trade-off between seeking permission and research taking place

The other split sample was asked a separate question in a similar format. This half of the sample was shown a description about how commercial organisations can access large volumes of hospital data to conduct research, where it is not possible to permission from patients each time it requires access to health data. Participants were assured that this data would not contain names or addresses, and were then asked which statement they agreed more with. Either:

- They would rather the NHS asks permission from patients when data is shared with commercial organisations, even if this means some research does not take place; or
- They would rather this research happens, even if in some cases permission is not asked.

Over half of respondents (54 per cent) would rather the NHS asks patients’ permission for research, compared with a third (34 per cent) who would rather the research happen even without permission.

Figure 6.12 – Support for research taking place using health data without permission being sought



Support for research without permission being sought

Which of the following statements comes closest to your view of commercial organisations seeking to access this kind of anonymised health data?*

A. I would rather the NHS ask patients’ permission to share anonymised data with commercial organisations, even if this means some of this kind of research does not take place

B. I would rather this research happen, even if in some cases the NHS does not ask for permission from patients

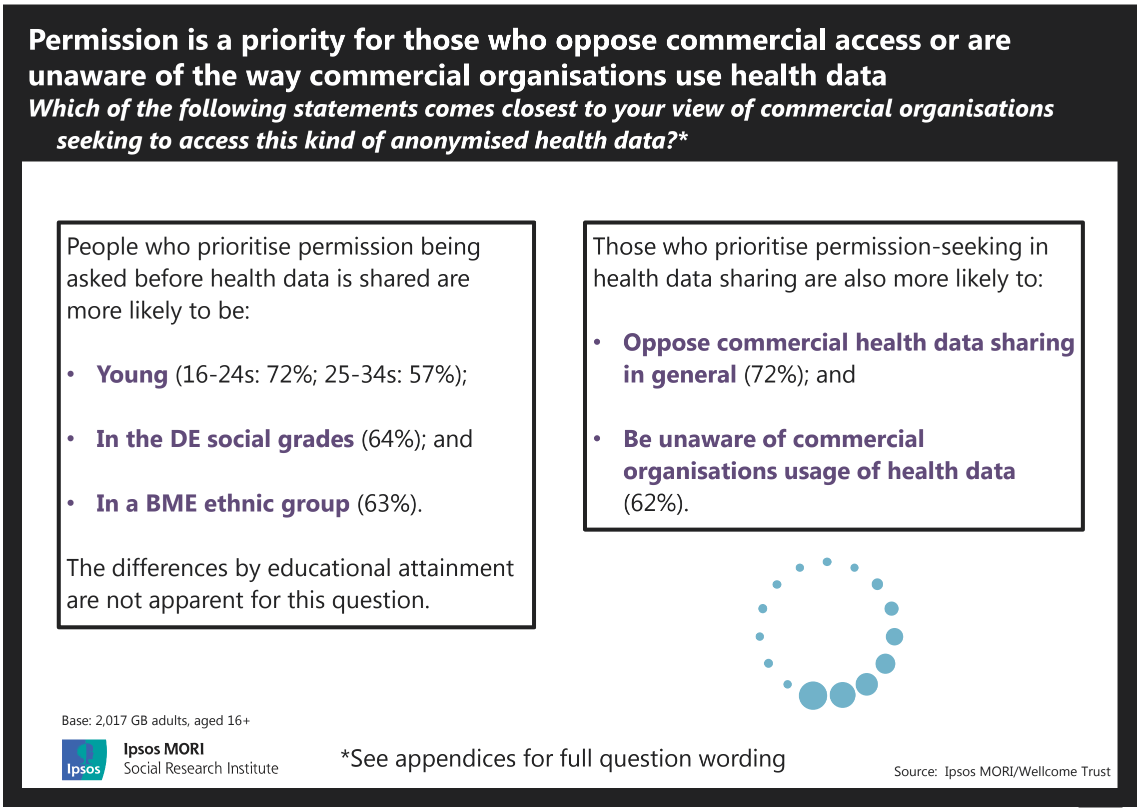
Base: 1,043 GB adults, aged 16+

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

Some of the characteristics of those who are most likely to opt for consent being sought above research taking place are shown in figure 6.13.

Figure 6.13 – Characteristics of those most likely to want consent to be sought



Permission is a priority for those who oppose commercial access or are unaware of the way commercial organisations use health data

Which of the following statements comes closest to your view of commercial organisations seeking to access this kind of anonymised health data?*

People who prioritise permission being asked before health data is shared are more likely to be:

- **Young** (16-24s: 72%; 25-34s: 57%);
- **In the DE social grades** (64%); and
- **In a BME ethnic group** (63%).

The differences by educational attainment are not apparent for this question.

Those who prioritise permission-seeking in health data sharing are also more likely to:

- **Oppose commercial health data sharing in general** (72%); and
- **Be unaware of commercial organisations usage of health data** (62%).

Base: 2,017 GB adults, aged 16+

Ipsos MORI Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

6.9 Safeguards

Respondents were randomly allocated to different split samples to test different safeguards. Each sample saw a different pair of statements and they were asked to agree with the one that came closest to their view. Agreement with statement A is agreement with commercial organisations accessing data with that particular safeguard in place, while agreement with statement B is agreement with 'We should not share data, as the risks to people's privacy outweigh the benefits'[18]. This question reflected a similar question in the 2014 study by Ipsos MORI and the Royal Statistical Society[19]. The aim of this split sample approach is to show participants only one option, to compare each safeguard separately, without concern that respondents may have been influenced in one way or another by a different safeguard. A sample presented with no safeguards was also put in place as a control.

As table 6.1 shows, the different safeguards had little impact on the acceptability of sharing health data with commercial organisations. There are no statistically significant differences between acceptance of health data being shared with an opt-out, anonymity, controlled

---

[18] Please see appendix for full question wording.

[19] Ipsos MORI, Royal Statistical Society (2014), 'New research finds data trust deficit with lessons for policymakers', https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-lessons-for-policymakers.aspx (accessed 06/12/15)

access or robust governance (described in the questionnaire as heavy fines and prison sentences) safeguard. Acceptability of commercial access to health data with a safeguard in place is fairly widespread, with between 56 per cent and 64 per cent agreeing with sharing data when a safeguard is in place. A large proportion of this agreement is also strongly voiced, with between 37 per cent and 44 per cent agreeing much more with commercial access to health data with a safeguard in place.

Table 6.1 – Agreement with data sharing with different safeguards in place

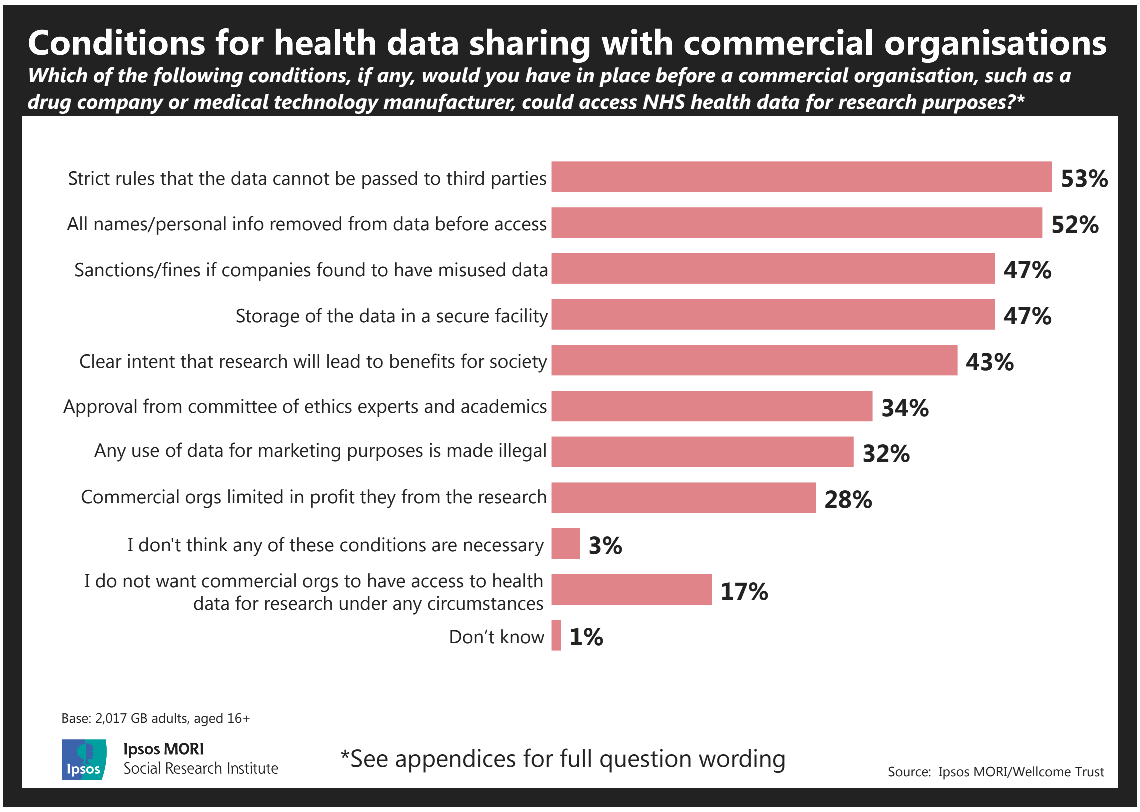
	Opt-out	Anonymity	Controlled access	Governance	NONE (control sample)
Base:	411	422	408	381	395
	%	%	%	%	%
Agree much more with A than with B **(Agree much more with sharing health data)**	41	43	44	37	30
Agree a little more with A than with B **(Agree a little more with sharing health data)**	20	14	20	24	19
Agree equally with both / don't agree with either	11	11	10	12	14
Agree a little more with B than with A **(Agree a little more with risks outweighing benefits)**	16	15	10	10	16
Agree much more with B than with A **(Agree much more with risks outweighing benefits)**	10	16	15	16	20
Don't know	2	2	2	2	2
% agree more with A than B	***61***	***56***	***64***	***60***	***49***
% agree more with B than A	***26***	***31***	***25***	***26***	***35***

Respondents are more concerned if this question is posed with no safeguard in place. The respondents that saw the control group statement – with no named safeguard – were less likely to agree with commercial access to health data (49 per cent) than where there was a safeguard in place. These findings reflect those from the Royal Statistical Society survey, and suggest that the precise nature of the safeguard is less important to improving acceptability than knowing that there are safeguards in place.

Acceptance within the different safeguards shows that some safeguards may matter to some groups more than others, but caution should be exercised as the sub-groups are particularly small due to the split-sample nature of the question. As well as this, the complex nature of the question means it is not possible to tell whether a sub-group difference is due to the safeguard in place, or due to the 'control' aspect of the question (i.e. general acceptance of health data sharing). Due to the small numbers here, we cannot analyse the differences between the sub-groups in relation to each safeguard with the sub-groups in the control sample, to see whether a difference is genuinely down to a safeguard making more of a difference to some groups than to others.

As well as this split sample experiment that explored the effects of different safeguards on responses, a more direct question was asked of respondents about the conditions they would like in place before commercial organisations could have access to NHS health data for research. Respondents were asked to choose the conditions they would have in place before commercial organisations could access health data for research.

Figure 6.14 – Conditions for health data sharing with commercial organisations



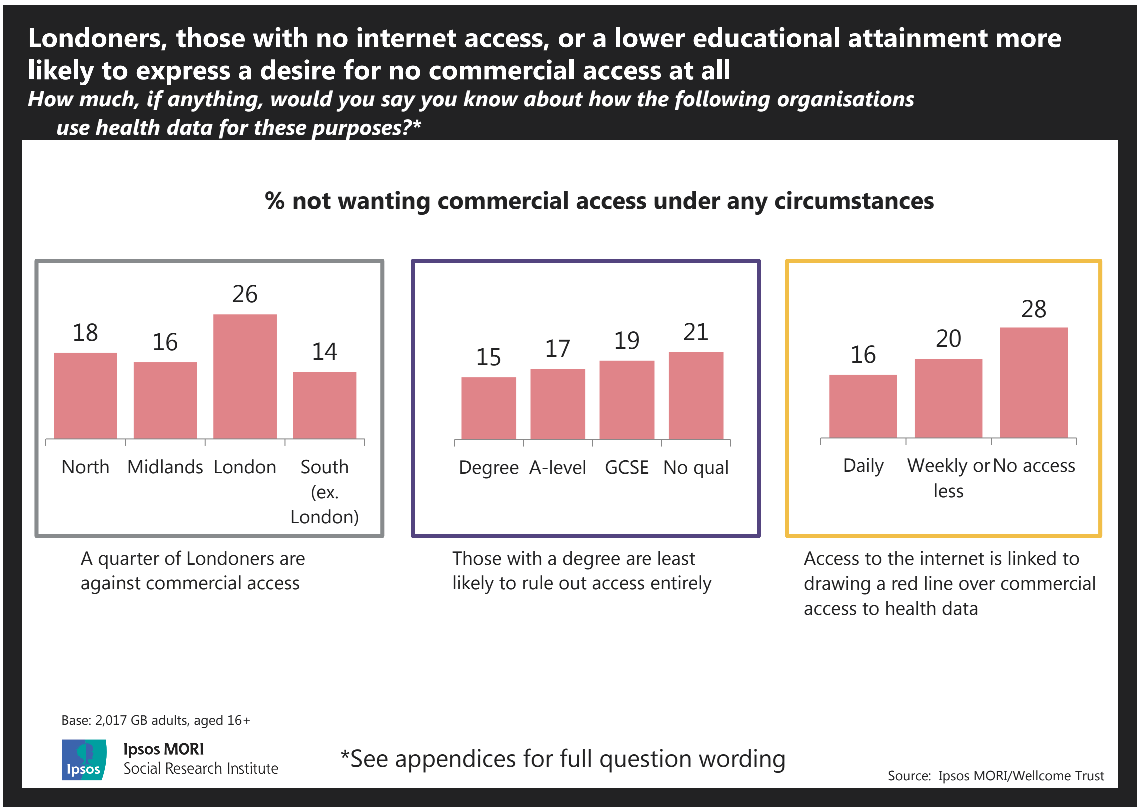
These conditions are clearly popular, and the top answer codes are all conditions that many people will opt for if given the choice. Over two-fifths of respondents mentioned three or more of these conditions (43 per cent). This suggests, firstly, that many people want several different conditions in place before commercial organisations have access to health data. While this seems to present a different picture of safeguards to what has been outlined at the beginning of this section, these findings do suggest that there is no individual safeguard that will act as a silver bullet to people's concerns. The broad mix of conditions named by respondents shows, again, that there is no specific safeguard that the public is waiting for, but more of a general expectation that policymakers are building in the procedures necessary to keep the public's data safe.

While there are many differences by sub-group, the differences in those who mentioned three or more different conditions stand out as of most interest for this report. Age, again, is a driver of the number of conditions mentioned, with under two-fifths of 18-24s and 25-34s mentioning this many conditions (38 per cent each), compared with a larger proportion of 45-54 year-olds (52 per cent) and 55-64 year-olds (48 per cent). This demonstrates, again, how much more comfortable older age groups feel in expressing their opinions on this subject matter.

Similarly, those with a degree are more likely to mention three or more conditions that they want in place (55 per cent), compared with under half of those with an A-level (45 per cent); two-fifths of those with a GCSE or equivalent (39 per cent), and 28 per cent of those with no qualification.

Respondents were also given the option of saying 'I do not want commercial organisations to access to health data for research under any circumstances'. Under a fifth opted for this response (17 per cent) and the composition of this group is depicted in figure 6.15.

Figure 6.15 – Characteristics of those most likely to express a desire for no commercial access at all



Londoners, those with no internet access, or a lower educational attainment more likely to express a desire for no commercial access at all

How much, if anything, would you say you know about how the following organisations use health data for these purposes?*

% not wanting commercial access under any circumstances

A quarter of Londoners are against commercial access

Those with a degree are least likely to rule out access entirely

Access to the internet is linked to drawing a red line over commercial access to health data

Base: 2,017 GB adults, aged 16+

*See appendices for full question wording

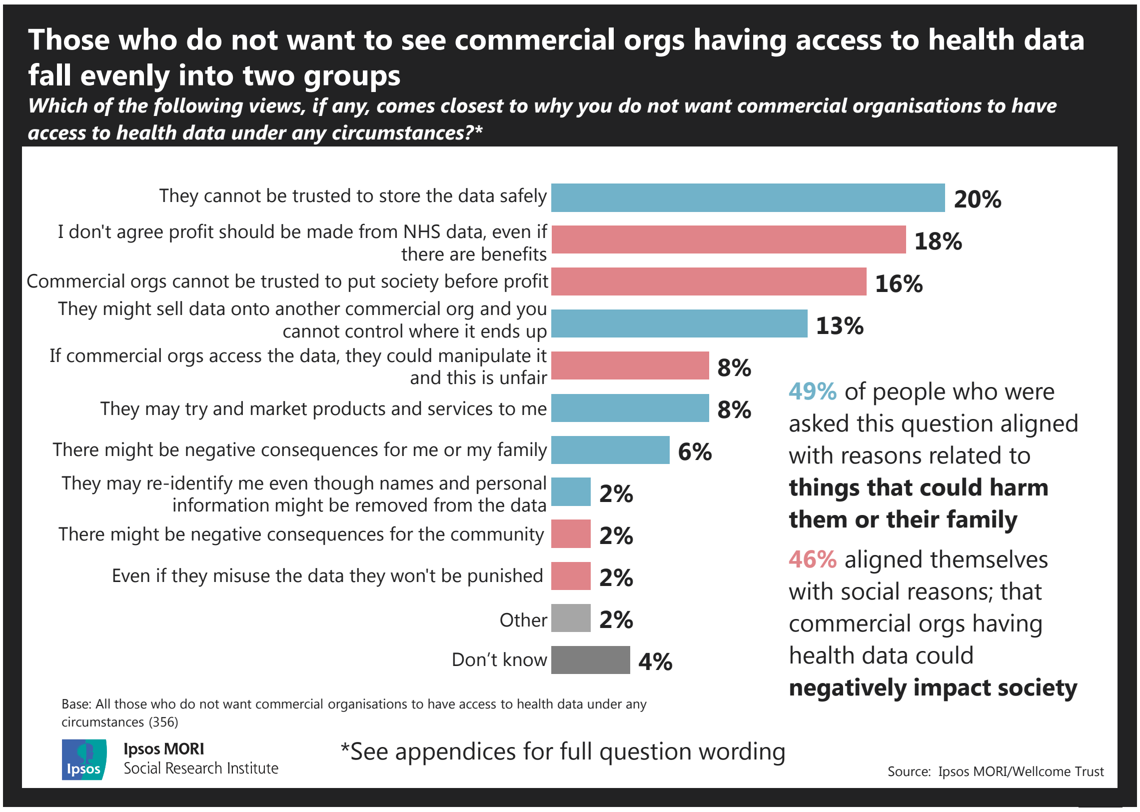
Source: Ipsos MORI/Wellcome Trust

The respondents who opposed sharing of any health data with commercial organisations were then asked a follow-up question, asking them which of a series of reasons comes closest to why they do not want to see commercial organisations have access to health data.

A fifth (20 per cent) of those who do not want commercial organisations to receive health data say that commercial organisations cannot be trusted to store the data safely. A similar proportion (18 per cent) agree most with the idea that profit should not be made from people's NHS health data – even if there are possible health and societal benefits as well. Sixteen per cent cannot trust commercial organisations to put the interests of society before profitmaking, while 13 per cent fear that they might sell the data on to another commercial organisation.

The findings for this question show that there are important themes running through the reasons why people are wary of data sharing with commercial organisations. The reasons that people most agree with fall into two main categories as shown in figure 6.16.

Figure 6.16 – Reasons behind the desire for commercial organisations not to have access to health data



Those who do not want to see commercial orgs having access to health data fall evenly into two groups

Which of the following views, if any, comes closest to why you do not want commercial organisations to have access to health data under any circumstances?*

49% of people who were asked this question aligned with reasons related to **things that could harm them or their family**

46% aligned themselves with social reasons; that commercial orgs having health data could **negatively impact society**

Base: All those who do not want commercial organisations to have access to health data under any circumstances (356)

Ipsos MORI Social Research Institute

*See appendices for full question wording

Source: Ipsos MORI/Wellcome Trust

Due to the small numbers of respondents being asked this question, there are no significant sub-group relationships useful to discuss.

7 Communicating with the public

7 Communicating with the public

7.1 Next steps: Pushing forward the debate on the role of data sharing in society

A more informed and engaged public could contribute better to the debate on the role of data sharing in the changing context of data, healthcare and society.

A discussion about data in society is very much like a discussion of science in society[20]. Data, like science itself, is not neutral, but is influenced by how society operates, and in turn is one of the things that drives the way society changes and behaves. How we legislate and regulate data use reflects the values of our society.

Potentially, the public's four questions (Why it is done, who is doing it, what data is accessed, how?) can be used as a start point or structure for debate on this.

Through the deliberative work several themes emerged. While participants did not necessarily articulate these ideas in exactly these terms, our analysis of what they did say suggests that these were the underlying ideas in play. Discussion of these will likely be core to any broader social discussion of the implications of commercial access to health data.

Important topics include:

7.1.1 What different kinds of value does data have?

Participants wanted companies accessing health data to demonstrate that they are making a contribution to public value and a real, long term benefit to health in the UK. They look to government and regulators to enforce it. In order to do this, we need a shared understanding of what that value actually is, and what kinds of public goods we want data analysis to create[21].

For the public in the workshops, the value of aggregate health data is in its long term value to society. This could be economic, or other sorts of value. They consider it this way for various reasons:

- In aggregate our data is a resource built up over years. Once a database has been 'sold off' it has gone, and the organisation who buys it pays once but can use it over and over

---

[20] This idea discussed in the Academy of Medical Science's recent paper which explores how a social contract for medical innovation might be brokered. http://www.acmedsci.ac.uk/policy/policy-projects/exploring-a-new-social-contract-for-medical-innovation/

[21] danah boyd & Kate Crawford (2012), 'Critical Questions for Big Data', Information, Communication & Society, 15:5, 662-679.

again to bring more benefit to themselves. The public want to ensure that ongoing value comes back to UK society not to private interest.

- Because health data is so personal, often based on someone's bad luck and ill health, and because individual-level data is captured in conditions of individual vulnerability and openness, (the service use mindset) participants are particularly keen that it should not be exploited for financial gain alone. They want it used to create social benefits for individuals and groups in a fair way.

- Participants in workshops felt that good use of data science should be used to support public goods such as the NHS. They do not want did not want anyone (public or private sector) to be able to co-opt health data for political ends, for example giving it to organisations who might have an interest in dismantling the NHS. For example, if work that the NHS could do is done instead by private companies who succeed because they have access to public data, this is felt to be wrong.

While aggregate data was thought of as a national resource, participants, especially those who did not think data had value, found it hard to conceive of who might own **individual** data. The financial implications of the future of healthcare are relevant here. Discussions touched on ideas of service delivery in future, for example if data is valuable, can those without money 'pay with data' for services? Are there unintended consequences or perverse incentives if this comes to pass, for example two-tier systems where the wealthier opt out, which might create bad data and potentially exploit vulnerable groups?

7.1.2 What should be the new social contract around health and digital data?

The basis of health data sharing is based on the Care Act 2014. Any data held by HSCIC can only be shared

- For the provision of health or adult social care

- For the promotion of health.

While this latter criterion suggests a public benefit, there is much discussion still to be had on the detail of what this really means when it comes to commercial access.

In the workshops there was a significant level of concern as to whether commercial interests will **prevent socially beneficial outcomes taking place** (with some mindsets very pessimistic about the future).

Resistance to insurance and marketing companies receiving health data was very strong. In the quantitative survey, even among those who support commercial access to health data more generally, nearly two thirds objected to health data being used to inform insurance pricing. More widely, insurance and marketing companies having access to health data was not supported. Data being a commodity at all was resisted in the workshops.

In the workshops participants said that this was because the purposes of insurance and marketing companies conflict directly with some perceived public interests. It is important to note, though, that while participants had a day to debate the issues, they were not pushed on the **social or economic value of insurance and marketing**. For example, a case could be made for economic benefits coming from these industries which might benefit society overall. Without evidence and prompts to frame this discussion, this debate was not well evidenced enough to take place in the workshops. There is scope for a further discussion or dialogue to dig deeper into this.

There is a debate beneath this about the role of state and commercial interests in healthcare at all, and the role of big data; a debate which is going on in wider society as well as in this social research study[22]. With the rise of the ability to collect and use data, what should change about the way health services are delivered? Participants touched on these issues, but there is much still to discuss. 'There is no job description for being a citizen', the Academy for Medical Sciences points out[23]; but there may be one in future, if sharing of data becomes a social right or duty rather than a choice.

Government, effectively, has stewardship of our aggregate health data, and is able to regulate access to it. The right deployment of 'the commons' has always been politicised, whether it be land or publicly owned infrastructure. The public consider the same discussions and tensions are present when we consider the 'national resource' of data. Participants felt that acting in the public interest here with this high-profile example of data sharing could be a key driver of public trust in government (or indeed, if handled badly it could severely erode public trust).

7.1.3 How can we situate conversations in the context of the future technology of healthcare?

While we discussed the idea of genetics and new ways of learning about individuals, plus touched on some ideas around personalised or predictive healthcare, the participants in the research did not know about the potential new innovations which could lead to both opportunities and risks for the healthcare system. The requirement for citizens to be empowered in order to take charge of their own health, alongside related issues such as how a population can all take advantage of online services equally, are concepts which need to be debated further.

In other research beyond this project, it is clear that people are not comfortable with new ways of collecting and sharing information in general. In Ipsos MORI & Demos' Wisdom of the Crowd study of ethics in using social media data for research, it was discovered that nearly three quarters (74%) would prefer to remain anonymous if a social media post was selected to be published in a research report; and over half (54%) agree that all social media accounts have

---

[22] These ideas are often reflected in emerging thinking for example in these articles:
https://www.insidehighered.com/blogs/library-babel-fish/negotiating-new-social-contract-digital-data
http://blog.politics.ox.ac.uk/social-contract-2-0-big-data-need-guarantee-privacy-civil-liberties/
https://www.bbvaopenmind.com/en/article/social-contract-2-0-big-data-need-guarantee-privacy-civil-liberties/
[23] http://www.acmedsci.ac.uk/policy/policy-projects/exploring-a-new-social-contract-for-medical-innovation/

the right to anonymity in social media research, even if the account is held by a public institution, private company or high profile individual.[24]

New technology, such as biometric devices and other wearables, might involve information captured passively to monitor individuals without them necessarily knowing this. This changes expectations of consent, and potentially increases 'context collapse' which will mean vulnerable people will need help to navigate the new terrain. The way that information is captured, especially in commercial contexts, can also potentially influence a discussion of data ownership.

Further investigation of the relationship of the public to the 'quantified self'[25] may be fruitful here. Participants in this project were uneasy with the idea of collecting information on their health and being accountable for their behaviours in order to receive services; they wanted to trust GPs to do it for them.

"My GP knows about my health, whatever I think I know on the internet."
Sutton Coldfield

A further observation is that across the workshops there was little awareness of **data science** and what it could do. Participants tended to focus on the political decisions which would come from outcomes of the analysis of health data. However, the way the data is structured and analysed creates further issues, relating to statistical bias, privacy, equity and governance[26]. There was very little awareness that, for example, the quality of data input into a system might lead to unrepresentative datasets which could lead to inaccurate evidence being provided. A discussion with the public on this matter would be fruitful.

Many of the topics above are already being investigated, for example by the Cabinet Office in a programme of research on data science; by the Royal Society investigating public views of Machine Learning; and by King's College and Ipsos MORI in a study of how the public responds to big data uses across a range of policy areas. There will be a need to draw together shared learning from all these projects as well as feeding into other policy strands such as the consultation of the Caldicott Review.

---

[24] https://www.ipsos-mori.com/Assets/Docs/Publications/im-demos-social-ethics-in-social-media-research.pdf

[25] The concept involves individuals producing, managing and using the streams of data they now create; discussed in this article http://www.economist.com/node/21548493

[26] 'Big Data for Policy Analysis: The Good, The Bad, and The Ugly', Schintler, Laurie A. and Rajendra Kulkarni, Review of Policy research, 31:4, 342-348.

8 Conclusions and recommendations

8 Conclusions and recommendations

The table below includes a summary of conclusions and relevant recommendations.

Table 8.1 – Conclusions and recommendations

Overall conclusions from qualitative and quantitative research	Recommendations for next steps
Most of the general public tend to accept commercial sharing of health data, as long as the four key tests are met.	If commercial sharing is on the table, the four key tests need to be applied upfront, so that the public can be reassured if they are asked to support this new way of using data.
Safeguards help the public feel reassured; the most convincing safeguards are those which regulate the profit motive in the interests of public benefit and create independent scrutiny and control. While safeguards on data handling are important, no one safeguard is a 'silver bullet'; the survey reveals that any safeguard is reassuring but no particular sanction or storage safeguard is more reassuring than any other.	Safeguards should be put in place first and be designed to enhance public control, for example opt-outs wherever possible, and overall designing a system where there are no hidden incentives for companies to behave badly.
There is a core group of those who do not want health data to be shared at all (17% do not want data shared for research under any circumstances; 25% would sometimes rather research did not go ahead than data is shared).	Policy and research interests are not likely to be able to sway this group. Opt-outs should be offered, along with clear communication about the safeguards and purposes of sharing, to minimise concerns among this group.
There are different views about different types of organisation, with some considered more acceptable than others for accessing data. • **Insurance is considered unacceptable** in the qualitative research and only 25% support it in the quantitative survey. • **Marketing was considered broadly unacceptable in the qualitative research, except in healthcare contexts,** and 38% supported this in the quant (NB a healthcare context was given). • **Third party access** to data was considered to be risky and not socially beneficial. The public do <u>not</u> want profit to be made from this resource without a company having a very explicit public benefit inherent in its work.	There is a need to identify and communicate a clear public benefit associated with data access if a company is involved, so that the public are reassured that profit motives will not override public benefits. To reassure the public, insurance and marketing uses of health data, should not be allowed. There is a need to restrict third party access and companies redeploying health data for further profit, and to open a broader debate about what value data has, and to whom should accrue that value.

Overall conclusions from qualitative and quantitative research	Recommendations for next steps
The public know little about some key areas: • Not aware of the current range of **uses of health data**, beyond in their own care • Underestimate the **amount of data** currently collected and used in healthcare • Do not understand why the NHS would need/want to allow commercial access to data, do not know **how the commercial sector contributes to healthcare currently** • Little understanding of the **status quo when it comes to safeguards** (some that participants want are already in place) • Confusion about the **specifics of data and data science in general.** (e.g. the difference between anonymised versus identifiable data, or definition of aggregate data).	These point towards a need to engage the public further, inform and communicate, in order to have a more informed social debate on the uses of health data. For example, there may be a need to inform the public as to how statistics work, what data is, and how it is combined into datasets, again in order that they can be informed on the subject. This will put the public in a better position to understand the real risks or benefits of data sharing.
There are many different mindsets and perspectives on commercial access to data; views of data sharing are influenced by opinions about society and commerce generally. New technology has given rise to new ways of collecting data, both actively and passively; and new ways of using the data to create knowledge. This has led to blurred lines between traditionally private and public sector ways of collecting data, causing a Context Collapse. The public are finding it difficult to navigate 'data contexts' as they shift and change, and are therefore very wary.	Codes of conduct may need to give explicit and separate consideration to the needs and fears of different groups of the public, for example ethical frameworks could be constructed which set out the spectrum of acceptability for different publics, as well as ensuring that all bases are covered for everyone. The public will need help to negotiate the context collapse, 'read' different data contexts more skilfully and feel more able to make decisions. NB: this will not necessarily lead to support for commercial access to data, but more information may mean members of the public feel better able to protect themselves.

Overall conclusions from qualitative and quantitative research	Recommendations for next steps
Overall, this project provides an evidence base to influence government and to use public views to improve the process of data sharing.	Policymakers and the research community should take into account the findings of this report when designing new processes and policies.
There is scope for a wider ongoing discussion. **A well-designed and timely discussion between policymakers, experts, research and clinical communities, and commercial organisations, as well as involving the public, could well shape the future of biomedical research, healthcare and notions of citizenship.**	

9 Appendix

9 Appendix

9.1 Qualitative research materials

9.1.1 Workshop/deliberative event discussion guide

To follow is the discussion guide used by the facilitators running the full-day deliberative events with members of the general public and the 'Children of the 90s' cohort members. This guide was slightly adapted for the shorter workshops with patients and healthcare professionals, although the broad topics covered and the majority of the exercises and prompts were the same. The guides for the Wrexham and Sutton Coldfield deliberative events were also slightly adapted to accommodate the improvisational actors' involvement. The adapted guides can be provided upon request (see the back cover of this report for research team contact details).

Please note that these documents were designed to be used as a **guide only**; although all of the key topics and exercises included in the discussion guides were covered with each group as far as possible, facilitators had the freedom to adapt the ordering and exact wording of the questions as best befitted the nature of the groups and the flow of the individual discussions they were moderating.

Discussion Guide for Access to Data workshops
General public workshops

This guide outlines the discussion that will take place between members of the public and facilitators at a series of dialogue workshops taking place around the country over the course of September and October 2015. The dialogue has been commissioned by the Wellcome Trust with the following objectives:

Project objectives

To identify factors that influence people’s attitudes towards commercial organisations accessing health, biomedical and genetic data; and to identify governance, safeguarding and communications actions that could help:

- improve trustworthiness of research uses and protections of data; and
- enable public trust in access to data to be developed over time.

Using case studies, in these events we will look at how acceptable different uses of data are, and how this varies by:-

- different types of data being used and accessed
- different types of users of data
- different purposes
- safeguards in place

9.40-10.00	Facilitators welcome observers, run through their roles and rules of engagement	Guidelines for observers/experts
10.00-10.30	Meet and greet participants, get coffee, photo permissions	
10.30-10.40 **10.40-10.50**	**<u>Welcome presentation and introduction</u>** **PLENARY:** *Presentation using slides* **TABLES (8-10 participants per table)** – Introduce self and note-taker and any observers at the table (participants free to ask questions at any time) – All views valid; please speak up and respond/agree/disagree to other points of view; try not to talk over one another; may need to interrupt to move discussion on. – Confidential with no direct attribution. – MRS Code of Conduct Permission to record – Housekeeping (phones on silent, location of toilets, any scheduled fire alarms, fire exits) – Introductions. Split into pairs for 5 min, then introduce partner to rest of table - **first name, where live, who with, what do with your days; last time you gave someone some data about yourself (not including just now when you signed in!)**	Set scene, housekeeping, allow participants to introduce themselves.
10.50-11.25	**<u>Warm-up discussion and brainstorming – what is known / believed already?</u>** **TABLES:** **A) What is data? POST-IT WALL EXERCISE** **Tell me everything that comes to mind when I say 'data'** – write down one thought per post-it (type of data, person/org asking for your data, possibilities of using data, risks) – anything that comes to mind *Participants stick post-its up on flipchart* DISCUSS RESPONSES **When do you share data?** GIVE EXAMPLES	This section will take participants through three areas: A) data generally B) medical data C) what people think the regulations are today around data. The point of this

	PROBE FOR PASSIVE COLLECTION VS. ACTIVE GIVING: – Direct to organisations – Buying products – Using services – public library, hospital visits, driving licence – Online/mobile behaviour – Amazon, Facebook, GoogleMaps **Who do you share information about yourself with?** SPONTANEOUS THEN PROBE – Companies – insurance, software developers, pharmaceutical – Government departments – NHS / GPs – Charities – Research bodies, universities – Partnerships of the above **What's the difference between these examples?** **How relevant is the type of data? Are you more / less comfortable with some than others? Why?** **How relevant is the organisation being given access to your data? Are you more / less comfortable with some than others? Why?** PROBE USING ABOVE LISTS **What kind of data are you thinking of here?** PROBE – Identifiable data (name, post code, date of birth, NHS number) – Coded but non-identifiable data (E.g. your store loyalty card number and shopping history but not your name and address) – Data that allows for demographic but not individual identification – Data collected at population level vs about you individually vs about a non-named person	section is to gauge their starting point: what they know or believe about existing commercial access to data (and other types of access). To this end we will be in listening mode, gathering views rather than providing information.

individually vs your behaviour vs your characteristics vs what happens to you…

How do you think the data is used? Why is it valuable to these organisations to have the data? FACILITATOR TO INTRODUCE THE IDEA THAT SOME DATA IS USED IN AGGREGATE: THAT THERE IS VALUE IN KNOWING ABOUT THE WIDER PICTURE E.G. FOR PUBLIC HEALTH OVERALL NOT JUST TO PROVIDE SERVICES TO INDIVIDUALS. THE NHS COLLETS DATA ON HOSPITAL WAITING TIMES, MADE UP OF THOUSANDS OF INSTANCES OF PEOPLE ATTENDING A&E.

B) Today we are mostly thinking about medical, health and genetic data.

Write on flipchart – medical data

What does your medical data mean to you?

PROBE

- difference between active data (where you consent to handover info about yourself)and passive data (where info about you/your behaviour is collected when you interact with a service)
- how collected
- by whom
- potential impact of sharing this information on individuals
- concerns over access (commercial/non)

Write on flipchart – health data

What does your health data mean to you?

PROBE

- difference between active and passive data
- how collected
- by whom
- potential impact of sharing this information on individuals
- concerns over access (commercial/non)

- difference between health and medical data (e.g. weight, calorie intake, exercise data)

Write on flipchart – genetic data
What does your <u>genetic</u> data mean to you?
PROBE

- how collected
- by whom
- potential impact of sharing this on individuals
- concerns over access (commercial/non)

Do you know what happens to these kinds of data? Beyond being used for your direct care? **Do you want to? What difference would it make?** PROBE anonymization, linkage, modelling, research

Let's talk about what makes data '<u>yours' OR OTHERWISE</u>

- What makes data 'yours'? (e.g. date of birth, identifiable details e.g. rare conditions, link to your name…) How about aggregated data of the level of your GP surgery, your postcode or local area – do you have the same views, same concerns, what is different?
- At what point does it stop being 'yours' and become general information about the population?
- Do you think you have a right to <u>opt out</u> of any data sharing, if the data is aggregate?
- How about if doing so affects the level of care or the consistency of care that is provided? By the NHS? By other organisations? To find out about public health across the country? Is there a 'social contract' in living here, is there data we *should* share for public benefit? EXAMPLE NHS collecting huge amount of data on waiting times

C) What safeguards or regulations do you think are in place around data uses?

SHOW FLOW CHART OF WHERE DATA IS COLLECTED IN (AND AROUND) THE NHS
What regulations do you think there are on how data can be used?

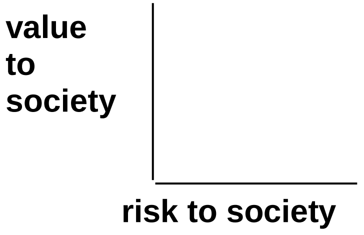
- On what data / level of detail
- In terms of who can use it

	– What is the “data protection act” and what does it cover? AFTER PROMPTING SHOW DESCRIPTION AND CHECK FOR ANY SURPRISES / MISCONCEPTIONS **D) In summary (if time)** What do you think are the main **<u>potential benefits</u> of giving access to your medical, health, genetic data? And <u>main risks</u>?** **How worried are you really? Where does this come from?** Personal experience, media, gut instinct *These are a first pass at the issues of the discussion. We will return to this throughout the discussion, at this stage we are trying to gauge participants’ openness to the overall concept, what do they spontaneously think of as ‘data’; how receptive they are, how many spontaneous worries they have (i.e. before being presented with detailed case studies).*	
11.25-12.30 **13-14 mins per case study**	**<u>Going through the case studies</u>** *IN TABLES* *Hand out individual participant packs including two slides per case study covering:* *1. What data is accessed* *2. How data is accessed and who has access* *Packs also include one pager on how healthcare data is collected currently, plus one on genetics to use with case study E* *Present each case study one at a time, check comprehension* *Around 12 mins per case study* *See order assigned to your workshop* *Ask key questions (below)* *Ask specific questions (on your slide pack)* *Compare between case studies as you go along*	To inform participants about existing case study examples and assess them in terms of acceptability/value /riskiness

	ASK FOR EACH:
	– **Spontaneous views / questions**
	– **How acceptable is it to share this type of data**
	– Where does this data come from? How acceptable is it to use it in this way? Identifiable with personal identifiers/coded with personal identifiers removed/aggregated? Allowing demographic but not individual identification
	– **How acceptable is it for this organisation to be involved – who stands to gain? Lose?**
	– **What do you think is the purpose of this** – and what else might happen if this happens?
	– **Views on how well data is protected/safeguards**
	– **Concerns** *(probe if necessary)*
	    – Personal privacy
	    – Data security/loss
	    – Unauthorised access
	    – Re-identification
	    – Discrimination/prejudice
	    – Lack of control
	    – Profit
	    – Data misuse
	
	*After showing the first 2 slides then move to **What Ifs.***
	NB the what ifs should be kept separate from the initial evaluation of how acceptable people find the current situation
	– **How acceptable is this?**
	– **Would this change your view on the whole case study? How?** *Probe extensively*
	
	Commercialisation/privatisation of the NHS may come up here as a concern. It is useful to get views on this but Wellcome Trust's main focus is on the public's concerns relating to research uses of data.	

12.30-	**LUNCH – please spend five minutes after having lunch writing up any questions and concerns**	

13.25	**that you have about what you've seen so far, on post-its and sticking them on flip chart page we have prepared.**	
13.25-13.45	**Finish off the case studies** *Final 2 case studies and summarise* **From what you have seen so far…** – **Which kinds of data are you most concerned about being accessed; and least concerned about; and WHY?** (both where it comes from e.g. GP records and whether identifiable/aggregated etc.); – **How about who accesses it; which companies or bodies are you most accepting of, least accepting of?** (e.g. NHS, Pharma, Industry, Gov) – **Which need the most safeguards or regulation?** – **Which are the best reasons for data to be shared? Which are less justifiable?** *Facilitator to collect key thoughts on flip chart, identifying red lines and tipping points* *After working through all case studies, participants fill out their* ***workshop questionnaire*** *individually, ranking each by acceptability*	Case study acceptability
13.45 – 14.00	**Overall evaluation of case studies - mapping** IN TABLE GROUPS On the wall we have prepared a large grid:	To explore other ways of assessing acceptability of the case studies

	 *Participants work together to map the case studies, sticking them onto the wall, trying to achieve consensus as to where they would fit* *Probe on why they make the choices they do, ask them to explain the thinking; what variables would need to change in order to make this work?*	
14.00-14.15	**Plenary summary of key thoughts and the overall evaluation** *Group returns to plenary and present back on their key thoughts so far, and where they placed the case studies on the map*	To draw out differences between groups
14.15-14.30	**AFTERNOON COFFEE BREAK**	
14.30 – 14.50	**Safeguards and trust** **Now we're going to think about suitable safeguards and rules for all this kind of activity.** **Thinking of everything we have looked at, ideally, what rules do you think should be in place? Why? FLIPCHART** **Which case studies / types of activity / organisations do these apply to? Why?**	To focus on safeguards and regulations more explicitly.

PROBE

- Consent – opt in/opt out? Should consent be sought for all uses of data beyond direct care?
- What if there was an oversight committee who made decisions about who should access data and under what circumstances, would you be happy to let them make decisions on your behalf? Or would you want to be asked each time?
- It’s not possible to consent to all aggregate data – now you’ve seen the case studies, what do you think about using aggregate data? And what about individual-level data that doesn’t have names/addresses attached? PROBE trade-off between slowing/preventing medical research vs. assurances of privacy?
- Ethics team to review new proposals (public input?)
- Audit trail in order to check who is accessing what and why
- Clear limits on how the data can be used
- Data destruction after use
- Secure storage (UK/EU/USA)
- Access to the results – public?

What difference does the type of organisation involved make? Why? What about collaborations between academics and industry?

Moderator note on consent: *there are different types of consent – ‘opt-in’ where individuals must actively agree to participate in a data-sharing exercise and ‘opt-out’ where people’s data are used by default unless they object. There are instances where health data is used and you have a right to raise an objection but that objection may be overruled. Explore pros and cons of different types of consent.*

Now thinking about the role of sanctions for misuse of data e.g. fines, suspension of licences, jail sentences
How important do you think these are? Crucial to acceptability of data-access or depends on the

<table>
<tr><td></td><td>scenario?
What kinds of sanctions do you think should be in place? WHY? PROBE fines, suspension of licences, jail sentences

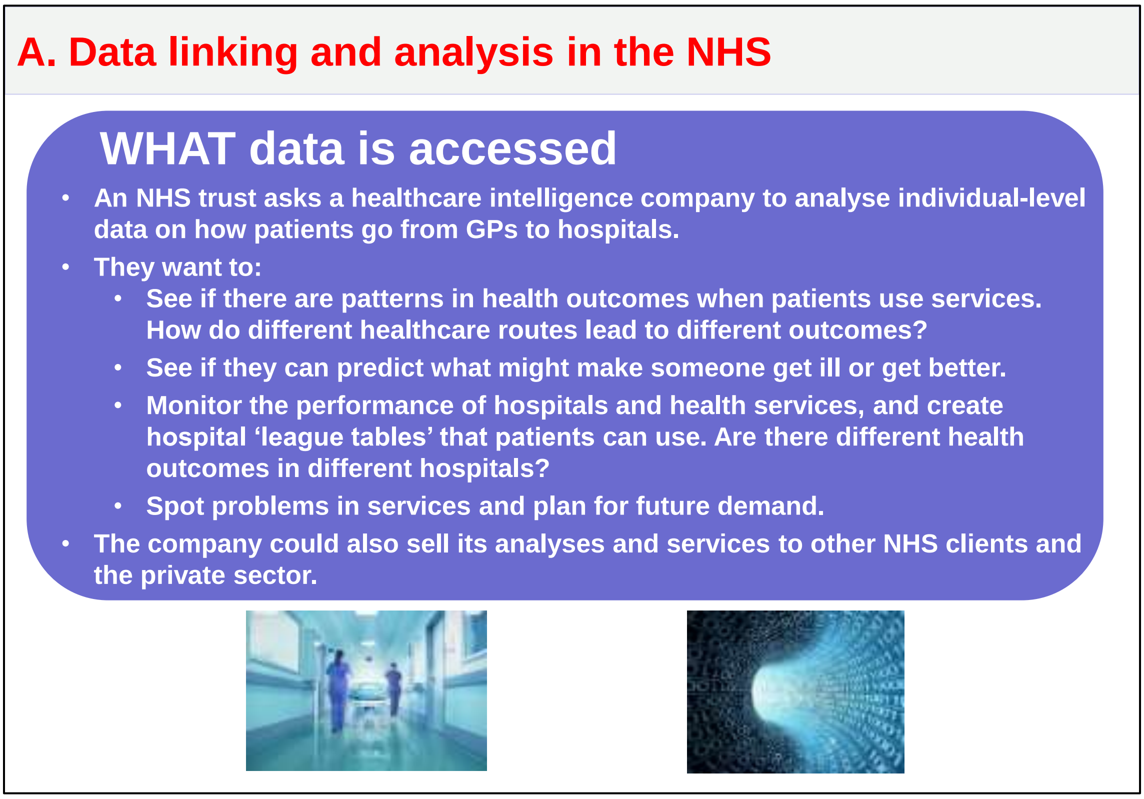
Who should be responsible for regulating this? Why?
PROBE
– Industry regulator
– National regulatory body
– International regulatory body</td><td></td></tr>
<tr><td>14.50-15.30</td><td><u>Principles of ‘good’ commercial access to health data</u>
IN TABLES
What do you think health services and commercial organisations need to consider to ensure they use health data in the best way for you – and for society as a whole?
Can you summarise some principles that should be in place
What advice would you give Wellcome Trust – they fund research and want their researchers to ask the right questions when they think about doing research involving health, medical, genetic data

Probes:
- Means of data collection
- Potential uses
- Whether future uses are known/unknown
- What kinds of organisation/collaborators
- Consent
- Timeframe
- Riskiness
- Public/personal benefit</td><td>Conclude by summarising key principles for ‘good’ commercial access to data and recommendations for Wellcome Trust</td></tr>
</table>

15.30-15.40	IF TIME/APPROPRIATE FOR THE GROUP: As well as thinking of these principles, I'd like you to think up a case study like the ones we've discussed that you would *like* to see happen – it has to involve a commercial company having access to your health, medical or genetic data *In plenary, each facilitator summarises principles from their group*	
15.40 – 16.00	**Summing up and close** *Thank participants for taking part today, and remind them how their contributions will be used (to help Wellcome Trust understand public views and inform future policy. Allow participants to make final points and ask final questions.* *Hand out incentives and sign form.* *THANK AND CLOSE*	To bring discussion to a close and administer incentives

9.1.2 Case study stimulus materials

To follow are each of the case study slides used in the discussion groups and workshops. Each case study slide is followed by a slide presenting a series of 'what if' scenarios; these were used following the initial discussions of the case studies themselves, in order to encourage further discussion of the key issues and drivers of acceptability and explore the nuances behind participants' initial reactions.



A. Data linking and analysis in the NHS

WHAT data is accessed

- An NHS trust asks a healthcare intelligence company to analyse individual-level data on how patients go from GPs to hospitals.
- They want to:
  - See if there are patterns in health outcomes when patients use services. How do different healthcare routes lead to different outcomes?
  - See if they can predict what might make someone get ill or get better.
  - Monitor the performance of hospitals and health services, and create hospital 'league tables' that patients can use. Are there different health outcomes in different hospitals?
  - Spot problems in services and plan for future demand.
- The company could also sell its analyses and services to other NHS clients and the private sector.

A. Data linking and analysis in the NHS

HOW is data accessed

- Individual-level data for more than 1 million patients is sent to the company by HSCIC, CPRD and the NHS trust. Data excludes names but includes information on visits to GPs, hospitals and social care services over a time period, e.g. 3 yrs.
- Individuals do not provide consent for data to be used.
- Data is stored in a 'cloud' database so that data from different parts of the health and care system can be compared and analysed.
- The company gathers data over time, creating a big database of health events.

WHO has access

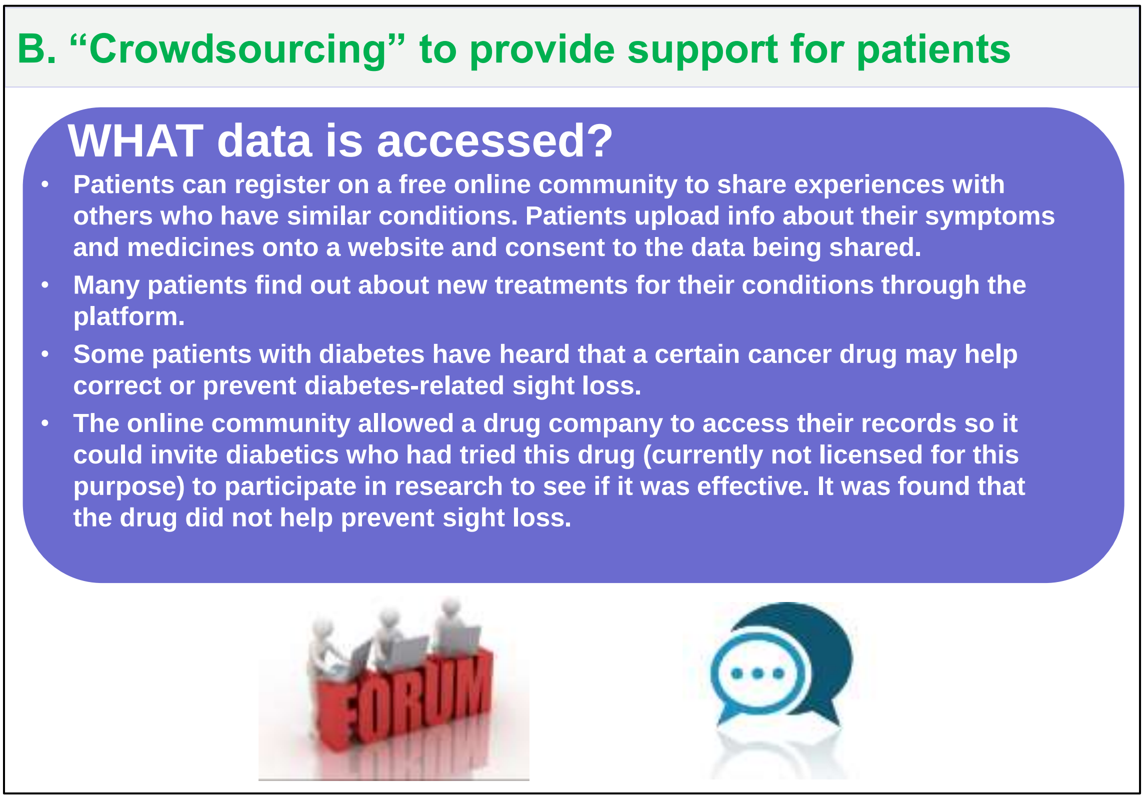
- Data analytics company hired by hospital trust
- Independent statisticians, academics
- NHS hospital trusts
- Private sector organisations who purchase the analyses

A. Data linking and analysis in the NHS

<u>What if we don't do this</u>: might we miss the next Mid-Staffs scandal? Patients died due to poor care discovered through this kind of comparison.

<u>What if:</u>

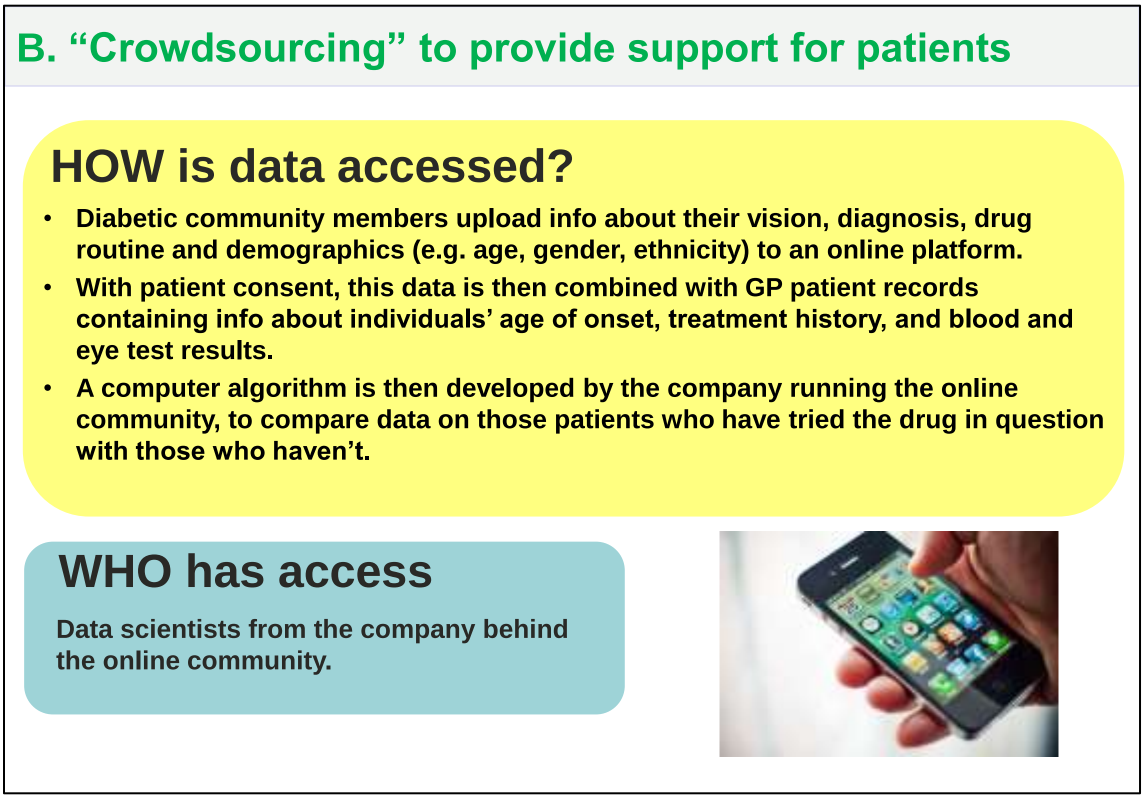
- **Government discovers certain areas of the UK have higher healthcare needs and services are cut to prompt 'behaviour change'.**
- **Your local hospital closes because it doesn't do as well as others.**
- **A private hospital provider predicts future healthcare needs across the UK and opens a private hospital near you.**
- **The analysis is sold to another company who links the data up with store loyalty card data, discovering where there is a market for a new weight loss company.**



B. “Crowdsourcing” to provide support for patients

WHAT data is accessed?

- Patients can register on a free online community to share experiences with others who have similar conditions. Patients upload info about their symptoms and medicines onto a website and consent to the data being shared.
- Many patients find out about new treatments for their conditions through the platform.
- Some patients with diabetes have heard that a certain cancer drug may help correct or prevent diabetes-related sight loss.
- The online community allowed a drug company to access their records so it could invite diabetics who had tried this drug (currently not licensed for this purpose) to participate in research to see if it was effective. It was found that the drug did not help prevent sight loss.



B. “Crowdsourcing” to provide support for patients

HOW is data accessed?

- Diabetic community members upload info about their vision, diagnosis, drug routine and demographics (e.g. age, gender, ethnicity) to an online platform.
- With patient consent, this data is then combined with GP patient records containing info about individuals’ age of onset, treatment history, and blood and eye test results.
- A computer algorithm is then developed by the company running the online community, to compare data on those patients who have tried the drug in question with those who haven’t.

WHO has access

Data scientists from the company behind the online community.

B. “Crowdsourcing” to provide support for patients

What if we don’t do this: cures for diseases are not crowd-sourced so it takes longer and is more expensive to find cures.

What if:

- **Patients download an app upgrade but don’t realise they’ve given automatic consent to data being sold to other companies.**
- **GPs are not in the loop – the patient community recommends a drug (marketed to them privately) but it is not suitable for everyone and someone dies due to lack of GP advice.**
- **Patients can sell the rights to their data to commercial companies interested in their condition. People living in poverty feel under pressure to sell their data.**

C. Monitoring safety of drugs and medicines

WHAT is being done

- **A public health regulator wanted to run an observational study to look at the treatment of high blood pressure using a specific drug.**
- **The drug was fully tested using all the usual clinical trials and classified fit for sale.**
- **To understand more about the potential side-effects of the drug, a long-term study of patients currently taking it was needed. For example, it was thought the drug might lead to increased risk of heart attack, stroke or death.**
- **The study showed these risks were no higher for patients taking the drug than those taking any other drugs for the treatment of high blood pressure.**
- **Thus the drug continues to be sold.**

C. Monitoring safety of drugs and medicines

HOW is this being done

- **Primary care data from CPRD was provided for over 12 million patients across the UK. The data came from around 650 GP surgeries signed up to the CPRD.**
- **Data included: age of patient, medical records, hospital admission and discharge dates. It did not include names, dates of birth or postcodes.**
- **Large numbers of patient data were needed for the results to be statistically reliable and to make useful comparisons across patients taking different types of drug.**
- **Patient consent was not given for GP records to be used.**
- **The public regulator used the data to compare the probability of heart attack, stroke and death in patients taking the drug with patients taking other treatments for high blood pressure.**

WHO is doing this

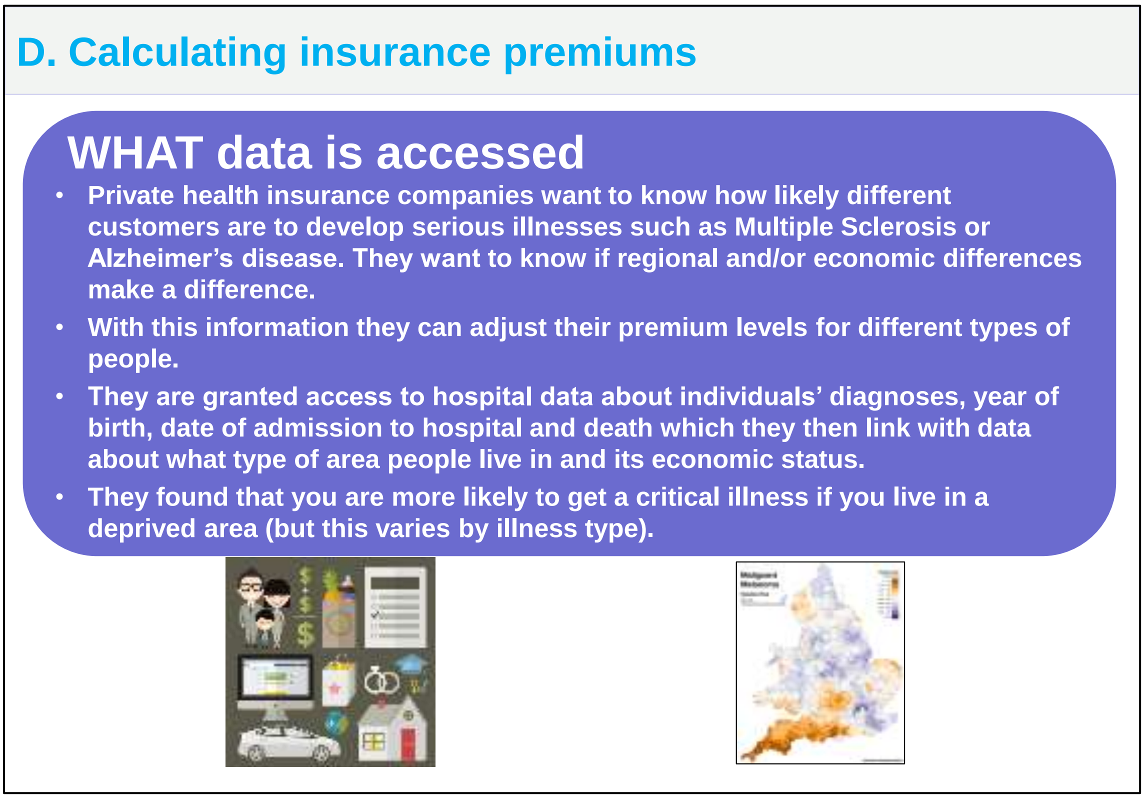
UK Public health regulator.

C. Monitoring safety of drugs and medicine

What if we don't do this: we might never find out that a popular drug is high risk or perfectly OK.

What if:

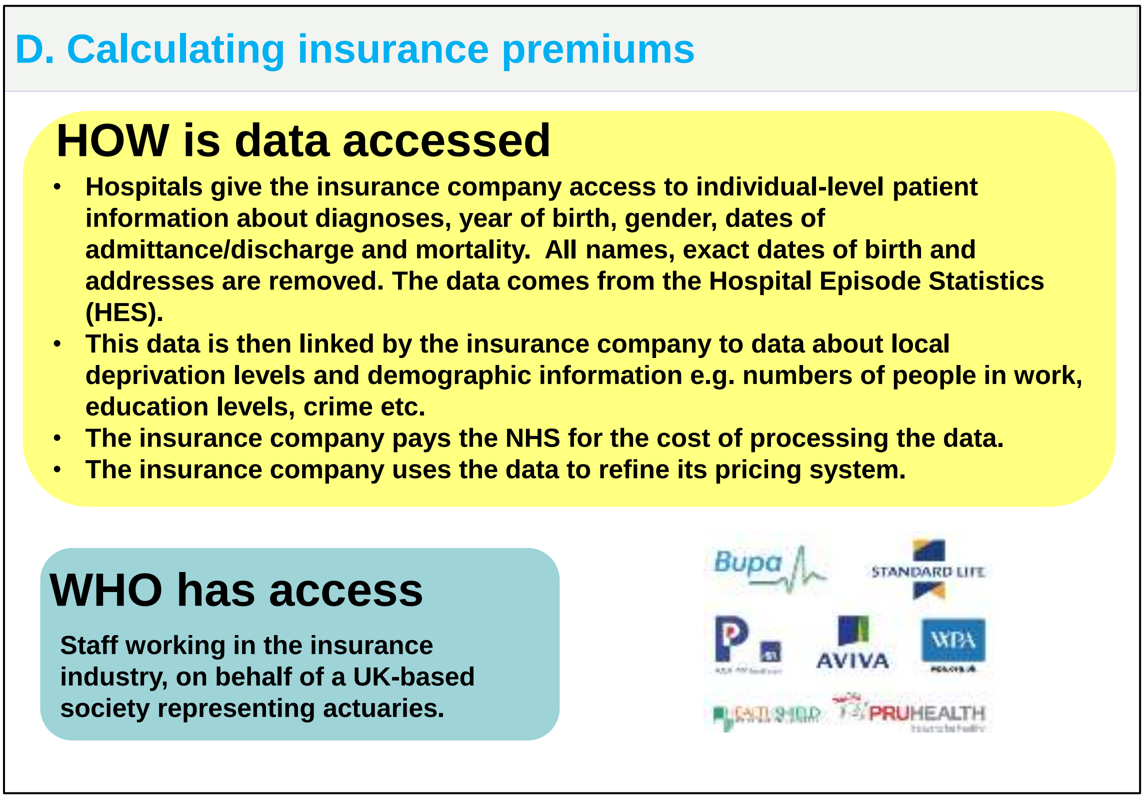
- **A drug company uses this data to identify gaps in the market and sell more costly medication to the NHS.**
- **The dataset is comparatively small and you can be identified – e.g. you're the only person in your GP practice who has been prescribed the drug in question.**
- **A drug company who manufactures the drug used to treat high blood pressure suspects that there are adverse risks and uses data from the study to protect itself from legal action. They use the data to deny patients compensation who believe they suffered poor heart conditions as a result of taking the drug.**
- **The data is shared with a US public regulator. UK patients are not informed.**



D. Calculating insurance premiums

WHAT data is accessed

- Private health insurance companies want to know how likely different customers are to develop serious illnesses such as Multiple Sclerosis or Alzheimer's disease. They want to know if regional and/or economic differences make a difference.
- With this information they can adjust their premium levels for different types of people.
- They are granted access to hospital data about individuals' diagnoses, year of birth, date of admission to hospital and death which they then link with data about what type of area people live in and its economic status.
- They found that you are more likely to get a critical illness if you live in a deprived area (but this varies by illness type).



D. Calculating insurance premiums

HOW is data accessed

- Hospitals give the insurance company access to individual-level patient information about diagnoses, year of birth, gender, dates of admittance/discharge and mortality. All names, exact dates of birth and addresses are removed. The data comes from the Hospital Episode Statistics (HES).
- This data is then linked by the insurance company to data about local deprivation levels and demographic information e.g. numbers of people in work, education levels, crime etc.
- The insurance company pays the NHS for the cost of processing the data.
- The insurance company uses the data to refine its pricing system.

WHO has access

Staff working in the insurance industry, on behalf of a UK-based society representing actuaries.

D. Calculating insurance premiums

What if we don’t do this: you end up paying very high premiums for insurance when there is little chance you’ll need cover.

What if:

- **The insurers find out that people living in social housing and on low incomes are more likely to develop critical illnesses and they market expensive insurance products to them.**
- **The company’s IT systems are hacked and hackers get access to a huge bulk of data about individuals’ diagnoses/year of birth/hospital admission dates.**
- **An NHS Trust carries out similar analysis to identify priority areas in need of certain healthcare services and end of life care at your local hospital drastically improves. (If we don’t do this: standards in palliative care are slow to improve.)**

What is… genetic data?

Every cell in your body contains a **full copy of all of your DNA**.

DNA is the **chemical code** that tells our bodies what substances to make in order to function and grow. It is divided up into **units called genes**.

A complete set of a person’s genes and all of the particles that store and give shape to their DNA is called **their genome**.

Think of your DNA as **lines in a library book,** each library book as one of your genes and your genome as the **whole library** containing all of the thousands of books that make up ‘you’.

You get half of your genes from your biological father and half from your mother. Everyone’s genome is unique, but there are many common characteristics that we share between us.

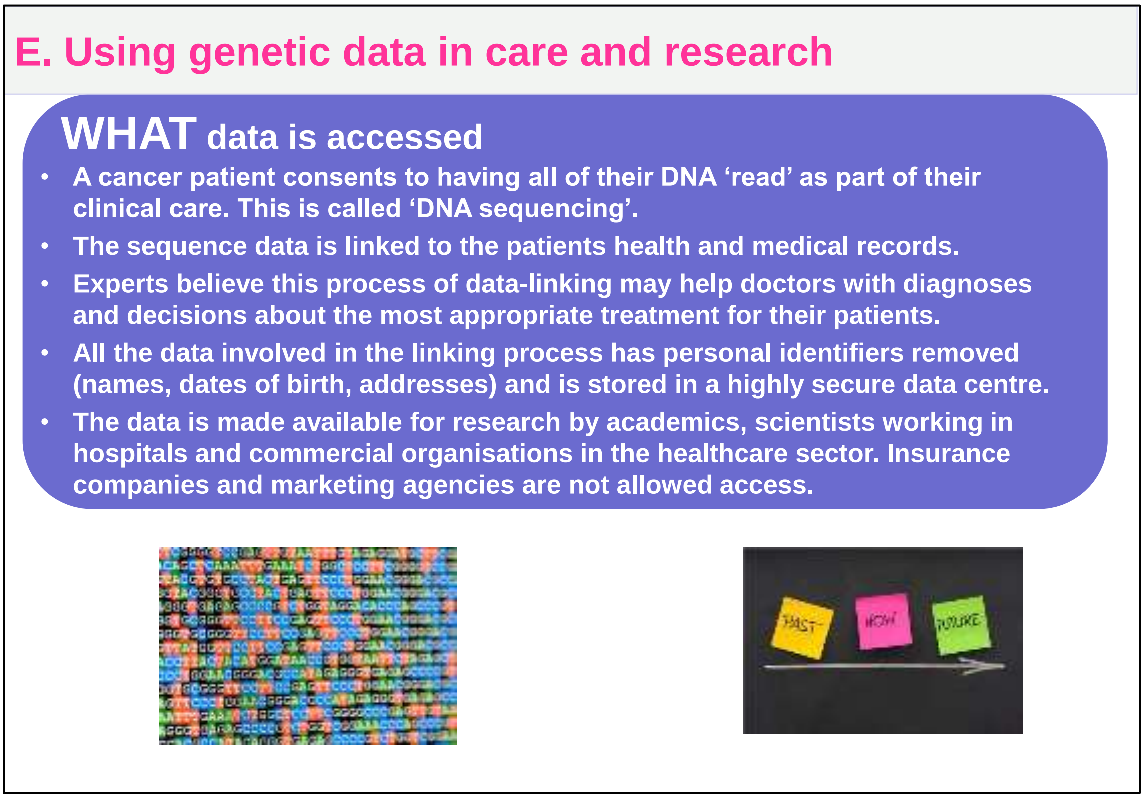
DNA

DNA makes up your…

Genes

Genes make up your…

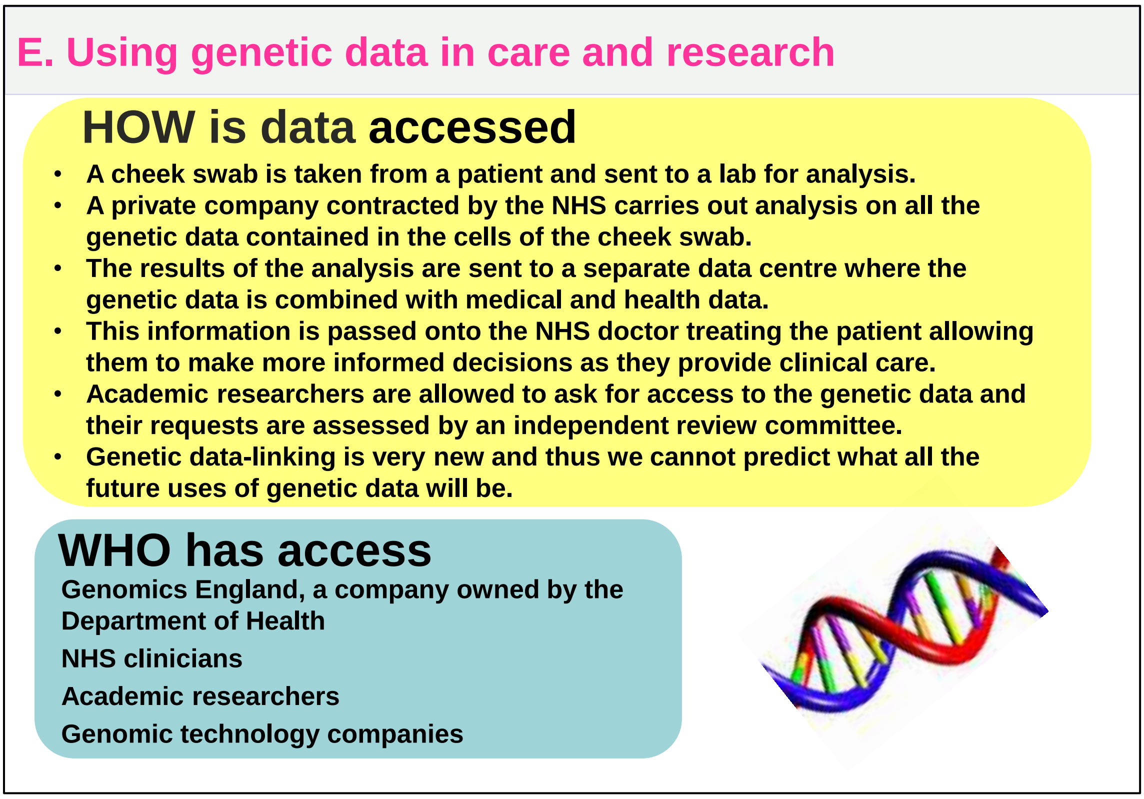
Genome



E. Using genetic data in care and research

WHAT data is accessed

- A cancer patient consents to having all of their DNA ‘read’ as part of their clinical care. This is called ‘DNA sequencing’.
- The sequence data is linked to the patients health and medical records.
- Experts believe this process of data-linking may help doctors with diagnoses and decisions about the most appropriate treatment for their patients.
- All the data involved in the linking process has personal identifiers removed (names, dates of birth, addresses) and is stored in a highly secure data centre.
- The data is made available for research by academics, scientists working in hospitals and commercial organisations in the healthcare sector. Insurance companies and marketing agencies are not allowed access.



E. Using genetic data in care and research

HOW is data accessed

- A cheek swab is taken from a patient and sent to a lab for analysis.
- A private company contracted by the NHS carries out analysis on all the genetic data contained in the cells of the cheek swab.
- The results of the analysis are sent to a separate data centre where the genetic data is combined with medical and health data.
- This information is passed onto the NHS doctor treating the patient allowing them to make more informed decisions as they provide clinical care.
- Academic researchers are allowed to ask for access to the genetic data and their requests are assessed by an independent review committee.
- Genetic data-linking is very new and thus we cannot predict what all the future uses of genetic data will be.

WHO has access

Genomics England, a company owned by the Department of Health

NHS clinicians

Academic researchers

Genomic technology companies

E. Using genetic data in care and research

What if we don't do this: doctors can't diagnose people quickly enough for effective treatment and patients die young.

What if:

- **A cancer patient consents to sharing genetic data because the are desperate to be involved in a trial of a new medication, as they want to find a cure. A decade later, experts find they can use that data to accurately predict life-expectancy. They contact the patient to see if they want to know.**
- **Government gives police access to a cancer patient's genetic data to help investigate a serious crime. The patient is found guilty of manslaughter and sent to prison.**
- **A private company buys the data storage centre from the government and sells access to the data to others for a profit.**
- **Analysis of genetic data means you can see what chance an unborn child has of developing a genetic condition. Abortion rates rise and commercial companies charge high fees to interpret genetic data.**

F. Pharmacists using Summary Care Records

WHAT data is accessed

- **The NHS wants all community pharmacists in the UK to have access to a summary of someone's health records and contact with the healthcare system (Summary Care Record, SCR). Pharmacists would have access to this summary when discussing prescriptions or giving customers health advice.**
- **The SCR contains information on a person's past prescriptions, allergies or bad reactions to medicine. 96% of the population have an SCR and patients can opt-out of having one via their GP.**
- **Allowing pharmacists to see patients' SCRs could:**
  - **Reduce the number of times pharmacists refer patients back to their GP or hospital;**
  - **Enable pharmacists to spot prescription errors.**

F. Pharmacists using Summary Care Records

HOW is data accessed

- **A pilot study was run involving 140 community pharmacies. Pharmacists approached patients and asked them to consent to having their SCR accessed.**
- **An advisory committee approved the study, including representatives from patient organisations and professional pharmacists.**
- **Only registered pharmacists were allowed to take part and they could only access the data through a highly secure network. Pharmacists were also required to log every time they accessed the network.**
- **If the data was used for anything other than direct patient care, the pharmacist would have his/her license withdrawn.**

WHO has access

A high street pharmacist such as Boots, Superdrug, Tesco pharmacy or Lloyds

F. Pharmacists using Summary Care Records

<u>What if we don't do this:</u> you have arthritis and are in excruciating pain but no medication at home. You go to your local pharmacist who you know and trust but they refer you back to your GP for a repeat prescription. You're frustrated and angry!

<u>What if:</u>

- **The NHS pays local pharmacists financial incentives to review what medication a person is using so many GPs stop doing this.**
- **An employee at a local pharmacy advises a patient suffering from high blood pressure on an appropriate dosage for their current prescription. They review the patient's SCR and see they've previously been prescribed medication for arthritis. The pharmacist sells them supplements for joint pain.**
- **Big pharmacies like Boots and Superdrug set official targets for selling non-prescription products to customers. Staff get financial bonuses as reward for successful sales.**

9.1.3 Case studies: detailed analysis of responses

The following table provides a summary of responses to each of the individual case studies and illustrates the factors influencing participants' judgement of their overall acceptability or unacceptability.

Data linking and analysis in the NHS

An NHS trust asks a healthcare intelligence company to analyse individual-level data on patient journeys, to see if there are different patterns in health outcomes, and predict drivers of service use.

Perceived value

- The **WHY** made this highly valuable; clear potential seen for patients and others; society to benefit, and this helped people overcome the risks.
- Helps ensure those delivering services are answerable to the public for efficiency.

  "League tables are not the whole story, you're just going by a snapshot, so this could be of use." General Public, Dundee
- Would help decision-makers with operational decisions.

Perceived risk

- Questions over whether the data really reflects what is going on – how reliable is it?
- Concerns over **WHY** - would the data be used to create legitimacy for unpopular political decisions?

  "The data may tell you that the hospital is not doing well – but what you do with that information is a political decision. You could close the hospital or you could invest in it more." General Public, London
- Concerns about **WHO**:
  - We could end up with hundreds of companies looking at our data and this could lead to unforeseen risks.
  - Don't always trust analysis companies/government to use the data wisely, they all may be subject to pressure to make money or repurpose the data beyond their original remit.
  - Data companies could also carry on gaining financial benefit from the data by repurposing it afterwards once they have it, which is seen as unfair as financial benefit should belong to all.
  - Private sector was a double edged sword – perceived to be efficient in the way the business is run, but motives are less trustworthy as profit is key concern. Participants frequently asked "couldn't the NHS do it?"
  - Also an opportunity cost to the public – could the public have gained more financial or other value from that data?

  "There's a certain level of accountability if it's being done by the government, but if it's a private company that is only accountable to its shareholders, profit, then the public good is undermined." Patients (rare conditions), Sheffield

Safeguards

- The **WHAT** was key: The data should not be identifiable and should be kept secure. This point was underlined by GPs with a sense of responsibility for the anonymity of the data they collect on their patients.

"One of the safeguards is anonymisation. Who has access to that data? It must be very high level access. You need to have a very secure system, I am not experienced in that but I suppose you can have rules about who specifically owns it in your service agreement. They may have to pay a license fee for example to access the data to benefit the NHS." GPs, Glasgow

- Not allowing the data to be sold on to other organisations beyond its original use was universally important.
- Consent became more important when private sector gets involved. HCPs felt they were guardians of the data and would need to ensure informed consent if it went to private sector, especially if it was to be repurposed more than once.

"It's not about worth, it's about ethical considerations. If you're going to use identifiable data you have to ask the patient." Hospital Doctors, Birmingham

- Need reassurance that data is stored in a way which preserves anonymisation.

Monitoring Safety of Drugs and Medicines

A public health regulator runs an observational study to look at long term side effects of a blood pressure drug. Primary care data from the CPRD is provided to compare the probability of serious adverse events compared to those on other drugs for high blood pressure.

Perceived value

- This was seen as valuable – for many it was a relief to know it was being done and further example of the solidity and trustworthiness of the public health system.
- The importance of the **WHY** outweighed the risks to any individual of a data leakage. These were usually thought to be minimal, in any case.

"They need to make sure the vaccines are safe, if it doesn't get back to you. You're just a number here, it's not a personal thing. It's only information required for the test" General Public, London

- This was seen as having a clear transactional value – public in aggregate need to be willing for this to take place, and if they are, they will reap the benefits.

Perceived risk

- Communication is key; in some workshops (e.g. the London pilot) there was a perception that the drugs/vaccines in the study were untested or not already known to be safe, which called the purpose of the study into question. This view changed when facilitators explained how population-wide studies took place in addition to drug development safety tests.
- Some wariness about **WHO** did the analysis. Considerable concerns that if the proprietary pharmaceutical companies did it, analysis would be skewed to make the drug look more acceptable. This related to a concern about **WHY** it was being done; some saw it as more risky if the work at all is done for profit, as this would increase the chance of bias towards a favourable outcome for the pharma company.

"The pharma company should pay for it, the regulator or academics should do it." General Public, Sutton Coldfield

- There was a general level of overall discomfort with money being made out of others' information; though it was not at its strongest when participants looked at this case study, it was still present.

Safeguards

- Peer reviewing the results, ensuring there is no bias involved. This was universal – almost all the groups felt it was important to have somebody independent involved to some extent. The pharmaceutical company should have to pay for independent assessment of the findings.

  "If you've got a drug company working alongside you, you want fantastic results. It has to be peer reviewed." General Public, Swansea

- Transparency was universally important.
- Views about consent were mixed. Participants had a natural desire for consent to be sought, especially if the drug was used in situations of serious or life threatening diseases and where people might have a strong personal reason to give or withhold it. However, groups did acknowledge that such consent might be impractical. Different groups came down on different sides of the debate eventually. (In Chapter 5 we discuss these different groups).
- Anonymization of results. The research was deemed to be safe because data would be very unlikely to be traced back to individuals, especially given the large sample size.
- Confining the data access only to its original purpose was important; perhaps by regulating the uses to which the pharma company could put it.

Using genetic data in care and research

Patients consent to having their genome sequenced as part of their clinical care. This is linked to their medical records to aid diagnosis and treatment, and is made available for research by academics, scientists and commercial organisations.

Perceived value

- This case study was perceived to be fascinating and the data was imagined to be very high value with huge potential benefits. The value was described in terms of the benefits to fuel medical research and drug developments and improve people's lives, (rather than financial value).
- Large social benefit in aiding diagnosis and helping people understand illnesses.

"There's a specific purpose, so most people would generally feel okay with it." General Public, London

- Patients with rare conditions show the most support as they have the most informed view of what the data could consist of and how it might be used.

  "I've actively consented to my genetic testing, for both of my conditions […] If they come up with something, not necessarily for me but for the future, I've consented to that – and if they find out things for me then that may be relevant for my son as well." Patients (rare conditions), Sheffield

Perceived risk

- All groups were very concerned that the future could hold great risks as well as benefits. There are uncertain implications, even those doing this research now don't know the full extent of what can be done with the data. This could include
  - Medical risks (what kinds of treatments might be possible?)
  - Ethical risks in terms of society and medicine (Who would reap the benefits? Would it be fair and equitable? Would it lead to eugenics, designer babies?)

  "This is companies playing God. It's disgraceful. Very unethical." General public, Dundee

  - Practical risks which could lead to significant social changes (Might the data be shared with other groups or agencies that could cause harm to individuals, for example law enforcement, employers?)
  - Practical implications for individuals if the data ended up with marketing companies (tailored, invasive sales approaches) and insurance companies (premiums being unfairly high to penalise the genetically disadvantaged, people not being able to access insurance all their lives because of genetic misfortune).

  "This is completely inappropriate, giving access to that [kind of] data. Until there is some evidence that this info will be of medical value, I can't see why they [non-medical companies with no public benefit] should have access." GPs, Glasgow

- Concerns around **WHAT** type of data is accessed – highly sensitive data, with extremely significant potential consequences if re-identified.

  "If they've got your genome on your database it doesn't matter if they haven't got your address and date of birth, they can still identify you." General Public, Swansea

- There was a generalised fear of a 'slippery slope' ; signing up now without knowing the implications of our decisions, as a society. Participants called for more discussion, expert advice and help before the public and government comes to final decisions about these types of data.

  "(Pharmaceutical companies) might come up with a drug where the results are 10% worse but makes them more money. If we're all going into this data, how can we ensure that the research is the best option?" Patients (severe conditions), London

Safeguards

- Consent was very important, however participants felt that at the moment it would be difficult for any citizen to give informed consent to due to unknown potential uses.

  "We need much more information on how this is used and the legal aspects of safeguarding." General public, Dundee

- Knowing **WHY** was a factor increasing acceptability – as was limiting the uses to which data could be put.
- **WHO** conducted the research and what this meant for transparency was key.

Pharmacists accessing Summary Care Records

The NHS wants all community pharmacists to have access to a summary care record. Pharmacists would have access to this with patient consent when discussing prescriptions.

Perceived value

- Makes service more efficient for all - relieves pressure on GPs – and for patients it's easier to go to the pharmacist than wait for a GP appointment.

"They know as much as the doctor does, and it's better than sitting in the cold miserable waiting room, waiting for your doctor." General Public, Belfast

- Benefits associated with **WHO** accesses the data. Pharmacists as 'another pair of eyes' could spot errors in prescriptions; people have positive experiences of this happening.

"I think it's important to educate people that they can go to a pharmacist and receive good quality information." Cohort members, Bristol

Perceived risk

- **HOW** the data is stored and accessed - Pharmacy setting seen as not very secure for the sharing of sensitive and personal data and participants felt this could lead to personal harms and lack of privacy.

"I've been in the phamacy and seen people getting methadone right in front of us [...] They're all addicted to drugs and you know - it's not very private." General Public, Belfast

- Worries about **WHY** it is being done – a slippery slope, part of the privatisation of the NHS/passing on responsibility that should be with GPs.

"My concern with this is not where it starts but where it ends up. You're diluting the job of the doctor to someone less qualified." Patients (severe conditions), London

- Some saw it as a marketing activity which would benefit the pharmacy and could harm the patient as it is a commercial transaction in an ostensibly 'public sector context' designed to sell products. Vulnerable people, who are unaware of the context, may be exploited (see Chapter 3 for a fuller discussion of how contexts around information are generally changing, making this a particularly pertinent concern today).
- Some have concerns about whether pharmacists are as educated/qualified as GPs.
- **WHO** – some felt more comfortable with this happening at a smaller, local pharmacy than a bigger organisation such as a pharmacy within a supermarket, seen as more commercially aggressive.

Safeguards

- Some would like to have the option to opt in, or opt out; are reassured that this is part of the plan.
- **WHO -** pharmacists/people accessing the SCR need to be appropriately qualified; people are not always sure of their level of education.
- Only the qualified pharmacist should be able to view people's SCR (i.e. if it's in a supermarket, other staff shouldn't be able to see it).
- In terms of **HOW** the data is accessed, the idea of a log of access is reassuring.
- Some feel that the patient, when giving consent, should have a say in **WHAT** information is shared, or that only information on the most recent ailment (for which the prescription

is needed) should be shared.

Calling Insurance Premiums

Calculating Insurance Premiums

Private health insurance companies use hospital data about diagnoses and hospital admissions and find that those living in deprived areas were more likely to develop certain critical illnesses.

Perceived value

- Very little to no perceived value; for most groups, negative value to society, because of total lack of trust that benefits would be public rather than solely to the insurance company. Perceptions of the industry are that the profit motive will always mean that the insurance company comes out on top.
- Did not see **WHY** this should be allowed.

  "Your premium could go down if they find out more detail about you? Hahaha! Like that would EVER happen." General Public, London.
- Some feel that it is fair for people at higher risk of illness to have higher insurance premiums, as long as this 'risk' judgement is based on fact; however, this is very much the minority view.
- The argument was made by facilitators that health insurance could benefit people by paying for their treatment and helping them insure against ill health, but this was seen as a very limited benefit.
- Some feel that if the NHS could make money from the information, this could be a potential benefit, but the risks are generally seen to outweigh this.

Perceived risk

- Unacceptable to most of the public because it's seen as unethical to access detailed health data for solely private gain with no link to improving public health.

"I can't see much benefit to the general public, but the only people going to benefit are the execs in insurance company. I'm not happy with it." General public, Dundee

- It is unethical to make premiums higher for those who are already unable/less able to pay.

  "What these companies are doing is making money, bracketing people, and then if you're in that box that's 'high risk' then they give you a higher premium." General public, Belfast
- Goes against the principles of the NHS, and universal healthcare. In the current system, insurance is seen as a 'nice to have' for the rich rather than an essential for everyone, so this data sharing is seen as inessential, socially destructive and potentially harming the NHS – which made the risks even more unjustifiable.

  "This goes against our society where we all pay for everyone's health. I'm proud of the UK for helping those who can't help themselves, this is being eroded." General public, Sheffield
- Judging, 'pigeon-holing' people based on where they live is also seen as unethical; participants accept in principle that it is possible to group people using statistics,

participants felt very uncomfortable with this being operationalised in any way that would harm or discriminate against an individual.

"It seems like it'll be high cost for certain people. It's morally wrong that they do it. So some people might be lucky…. but not me [with] my job…low paid job [will I lose out?]." General public, Dundee

- Assumptions that insurance companies would not put as much emphasis on data security (reflecting generally low levels of trust, and memories of commercial data losses of the past from the private sector).

  "The person at the insurance company isn't going to be as strict on the data protection act." Cohort members, Bristol

Safeguards

- Some would be reassured if the data was anonymous and could not be traced back to the individual; however, most still feel that anonymous and aggregate data can be manipulated or used to the detriment of groups of people, regardless of **WHAT** data is accessed.
- People would want to be informed, and for the data to be publicly available so they could see what data was being used for this. If the project was to go ahead at all (which most people did not want) participants would require the company to undertake a very well-communicated and well-designed option to opt out.
- There was no perceived difference between actuaries working for insurance, and insurance companies themselves – so differentiating access on this basis did not work as a safeguard.

Crowdsourcing to provide support for patients

Patients register on a free online community to share experiences and symptoms. The online community allows a drug company to invite diabetics to participate in research into the efficacy of a drug to treat sight loss.

Perceived value

- Was seen as a positive and helpful resource for patients to be able to share experiences and get support.
- Could potentially help improve treatments, and is a relatively low-cost and potentially far-reaching source of information with which to do so.

  "The public themselves are engaging and initiating - it's a cheaper way of doing it…turning the business of research on its head." General Public, Dundee

  "I like the idea of crowdsourcing because you put experts in with a group of people who are easy to access and willingly giving their data. By putting specialists in you get a benefit out of it." General Public, London

Perceived risk

- Concerns about the accuracy of **WHAT** information people are sharing – worries about people incorrectly 'self-diagnosing', and about how rigorous those **WHO** run the site would be.

  "You would wonder how academic and accurate they could be, and what kind of people could be running that" General Public, Dundee

  "It depends on the philosophy behind the group owners […] Is it for personal gain, for

the company to make millions of dollars? Is it for the help of another human being? Because I'm ok if it is. But I'm not ok if there's corruption behind it." General Public, Belfast

- People want to see what the benefits are, and that the data is being used in a way that clearly benefits the community whose data is being used; there were worries that this might not be the case.
- Data being sold on to marketing and insurance companies.
- Major concerns among HCPs about the rigorousness of the research process and value of outcomes

"If it turned out the drug was harmful was that going to be swept under the carpet." GPs, Glasgow

Safeguards

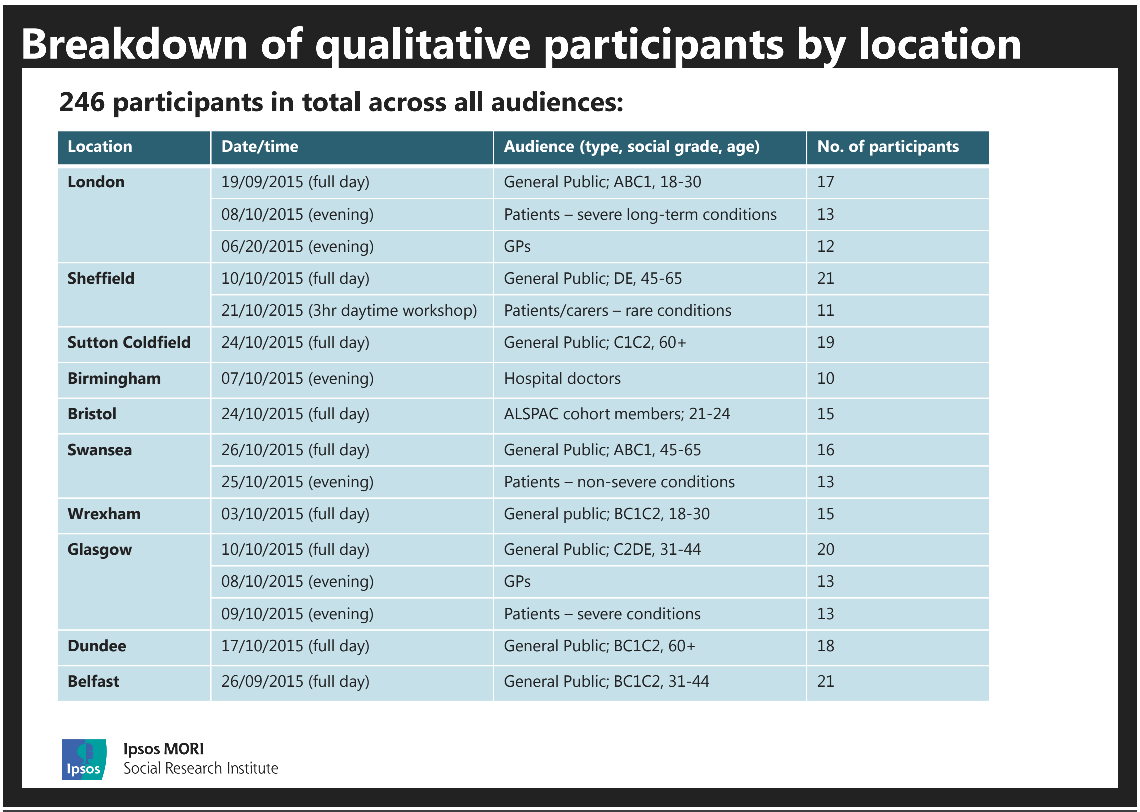
- Participants felt this would be an opt-in system by its nature; you can choose to participate in it, or not. But: needs to be made clear exactly **WHAT** you are signing up for. Data extracted from it should be anonymised.
- **WHO** – site needs to be administered carefully, to militate against misinformation, inaccurate data and false advice. Some participants felt more confident when given the different variable as an example that the site could be run by the NHS, or the NHS could oversee it in some way. There was a strong feeling that GPs/health professionals should be involved. But HCPs questioned the value of 'real-world data' and the biases inherent in any self-selecting sample.

"It's called patient testimonies, it is rubbish. The selection bias in this example is phenomenal. It would never be published in a journal. What kind of people are going to go on to this website? People who are dissatisfied, are a bit eccentric perhaps. You will not get a cross section of the population." GPs, Glasgow

- Should only be health-related organisations accessing the data, and there need to be strict regulations on who it can be used by.
- **HOW** – data management should be properly regulated by an independent body.

9.2 Qualitative fieldwork; detailed breakdown of workshop locations and participant profiles

The chart below shows a detailed breakdown of workshop locations and participant profiles, including the approximate social grade and age profiles of the general public groups.



Breakdown of qualitative participants by location

246 participants in total across all audiences:

Location	Date/time	Audience (type, social grade, age)	No. of participants
London	19/09/2015 (full day)	General Public; ABC1, 18-30	17
	08/10/2015 (evening)	Patients – severe long-term conditions	13
	06/20/2015 (evening)	GPs	12
Sheffield	10/10/2015 (full day)	General Public; DE, 45-65	21
	21/10/2015 (3hr daytime workshop)	Patients/carers – rare conditions	11
Sutton Coldfield	24/10/2015 (full day)	General Public; C1C2, 60+	19
Birmingham	07/10/2015 (evening)	Hospital doctors	10
Bristol	24/10/2015 (full day)	ALSPAC cohort members; 21-24	15
Swansea	26/10/2015 (full day)	General Public; ABC1, 45-65	16
	25/10/2015 (evening)	Patients – non-severe conditions	13
Wrexham	03/10/2015 (full day)	General public; BC1C2, 18-30	15
Glasgow	10/10/2015 (full day)	General Public; C2DE, 31-44	20
	08/10/2015 (evening)	GPs	13
	09/10/2015 (evening)	Patients – severe conditions	13
Dundee	17/10/2015 (full day)	General Public; BC1C2, 60+	18
Belfast	26/09/2015 (full day)	General Public; BC1C2, 31-44	21

Ipsos MORI
Social Research Institute

9.3 Quantitative methodology and topline survey results

9.3.1 Publication of the data

Our standard Terms and Conditions apply to this research, as to all studies we carry out. Compliance with the MRS Code of Conduct and our clearing is necessary of any copy or data for publication, web-siting or press releases which contain any data derived from Ipsos MORI research. This is to protect our client's reputation and integrity as much as our own. We recognise that it is in no-one's best interests to have research findings published which could be misinterpreted, or could appear to be inaccurately or misleadingly represented.

9.3.2 Capibus methodology

Capibus provides a high quality sample of adults aged 16+, representative of the population at a national and regional level. Capibus uses a two stage random location design to select respondents to take part in the weekly survey. The two stages are as follows:

Stage One - Selection of Primary Sampling Units:

A total of 154-180 Postcode Districts (e.g. HP3) are selected from across Great Britain, and this selection is stratified by region so as to ensure balance and representativeness at the high geographic level.

Stage Two - Selection of Secondary Sampling Units

At this stage, one paired Census ONS Output Area is randomly selected from each Postcode District. This then becomes the secondary sampling unit.

An Output Area (OA) is a very small area made up of about 125 addresses. Ipsos MORI carried out a detailed and scientific pairing exercise following the dissemination of the Census OA boundaries in order to pair-up OAs with their nearest neighbours to form Double Output Areas. These Double Output Areas, which contain an average of 250 addresses are used for most of our residential public opinion work where a large number of interviews are required in each point. Double Output Areas thus constitute the secondary sampling units for Capibus.

Double Output Areas are selected with probability proportional to their size. This gives the larger (more populous) Double Output Areas the greatest chance of selection. The purpose of this is to ensure that each address will have the same chance of being selected in the survey, regardless of the size of the Double Output Areas in which it falls. It is a standard technique that Research agencies adopt where a fixed number of interviews are needed in each sampling point.

A sophisticated set of sampling programs run to ensure that the collective selection of the Double Output Areas in any given week is balanced and match the demographic profile of the population. This is done through balancing the sample across the various CACI ACORN groups (which themselves cover a number of resident characteristics), ethnic profile and rurality. Half of the sampling points are drawn initially - their collective ACORN profile is then checked and compared with the national population. In the remaining half, points from the thus-far under-represented ACORN groups are then sampled with appropriately increased probability of selection.

Adopting this approach helps to eliminate any possible bias in the sample caused by interviewing people all with the same background. Using ACORN allows us to select OAs with differing profiles such that we can be sure we are interviewing a broad cross-section of the public, since clearly even people of the same age and working status may have a different viewpoint depending on their background. Because the sampling process is repeated every week, the Capibus sample is matched wave on wave, making it ideal for taking successive measurements on the same issue.

Stage Three - Selection of addresses

Within the 154-180 Double Output Areas which have been, each week, selected completely at random, we set our interviewer quotas on gender, age, working status and social grade; the numbers of each group we interview will depend on its exact local Census population. The total number of interviews per Double Output Areas is between154-180. This is a common approach for ensuring that a sample is nationally representative. Although the final stage of Capibus is quota-based, the sampling scheme for this survey overall is predominantly random. The sample selection down to the small Double Output Areas geographies is random.

Geography

The regions, which were defined at county level by the former Department of the Environment, are broken down as follows:

Total % of Sample	Region
5%	Wales
6%	North East
11%	North West
9%	Yorkshire
9%	West Midlands
7%	East Midlands
4%	East Anglia
8%	South West
19%	South East
12%	London
9%	Scotland

Fieldwork

Fieldwork is carried out by Ipsos MORI using CAPI (Computer Assisted Personal Interviewing). All interviews are conducted face to face, in the home – one interview per household. No incentives are offered to respondents.

Weighting and data processing

Data entry and analysis are carried out by Ipsos MORI data processors. All information collected on Capibus is weighted to correct for any minor deficiencies or bias in the sample. Capibus uses a 'rim weighting' system which weights to the latest set of mid-year census estimates and NRS defined profiles for age, social grade, region and working status - within gender. The idea of rim weighting is to provide the 'best weighting', or least distorting, by using computing power to run a large number of solutions from which the best is chosen. Thus 'Rim weighting' is superior to the more common system of 'Cell weighting'. Additional profiles used include Tenure and car in household for example.

Statistical reliability

Because a sample, rather than the entire population, was interviewed for this survey, the percentage results are subject to sampling tolerances – which vary with the size of the sample and the percentage figure concerned. For example, for a question where 50% of the people in a (weighted) sample of c. 2000 respond with a particular answer, the chances are 95 in 100 that this result would not vary more than two percentage points, plus or minus, from the result that would have been obtained from the census of the entire population (using the same procedures). An indication of approximate sampling tolerances is given below.

Approximate sampling tolerances applicable to percentages at or near these levels (at the 95% confidence level)			
Size of sample or sub-group on which survey result is based	**10% or 90% ±**	**30% or 70% ±**	**50% ±**
2,017 (All people interviewed)	1.3	2.0	2.2

Strictly speaking the tolerances shown here apply only to perfect random samples - in practice good quality quota sampling has been found to be as accurate - although clustering of the sampling points and non-response weighting can have an influence on statistical reliability.

9.3.3 Quantitative survey topline results

Technical details

Ipsos MORI interviewed a representative quota sample of 2,017 adults across Great Britain aged 16 and over. Face-to-face interviews were conducted in-home between 30 November and 11 December 2015. Data is weighted to the known population profile of GB.

An '*' indicates a finding of less than 0.5%, but greater than zero.
Where percentages do not add up to exactly 100% this is due to computer rounding, the exclusion of "don't knows" or to multiple answers.

Q1. **Health data collected from patients in hospitals and GP practices can also be used for research into diseases and treatments, and for planning healthcare services. When used in this way, health data has personal information removed, such as patients' name and address.**
How much, if anything, would you say you know about how the following organisations use health data for these purposes?

	The NHS %	Commercial organisations, such as drug companies and medical equipment manufacturers %	Academic researchers %
Base:	*All respondents (2,017)*	*All respondents (2,017)*	*All respondents (2,017)*
A great deal	12	5	5
A fair amount	21	11	13
Just a little	29	25	25
Heard of, know nothing about	21	27	25
Never heard of	16	31	31
Don't know	1	1	1
At least 'heard of'	83	68	68

Q2. **As you may know, the NHS and other health services collect data about individuals, for example GP records and hospital visits. This data is used for the direct care of patients, but it can also be useful for hospitals, commercial organisations, researchers and others. Allowing access to data can bring benefits, such as enabling research to find more effective medical treatments for example. However, some people worry that allowing others access to health data will be a risk to their privacy and security, for example if it is possible to re-identify individuals.**

Overall, which of the following statements is closest to your view[27]?

	SPLIT SAMPLE A – Opting Out	SPLIT SAMPLE B - Anonymity	SPLIT SAMPLE C – Controlled Access	SPLIT SAMPLE D – Governance/ Enforcement	SPLIT SAMPLE E – Control question
Base:	*(411)*	*(422)*	*(408)*	*(381)*	*(395)*
	%	%	%	%	%
Agree much more with A than with B	41	43	44	37	30
Agree a little more with A than with B	20	14	20	24	19
Agree equally with both / don’t agree with either	11	11	10	12	14
Agree a little more with B than with A	16	15	10	10	16
Agree much more with B than with A	10	16	15	16	20
Don’t know	2	2	2	2	2
Agree more with A	**61**	**56**	**64**	**60**	**49**
Agree more with B	**26**	**31**	**25**	**26**	**35**

---

[27] For full statement wording, please see page 152.

Q3. **To what extent, if at all, would you support your health data being accessed by commercial organisations if they are undertaking health research?**
By health data, we mean information that is routinely gathered about individuals by the NHS when they receive care. The information given to the organisation would NOT include your name and address or any contact details.

Base:	*All respondents (2,017)* %
Strongly support	18
Tend to support	35
Neither support nor oppose	19
Tend to oppose	13
Strongly oppose	13
Don't know	2
Support	54
Oppose	26

Q4. **To what extent do you agree or disagree with the following statements?**

	"My health data currently has financial value to others in that it can be used to save or make them money."	"My health data currently has a value to society in that it can be used to help improve things for people other than me."
Base:	*All respondents (2,017)* %	*All respondents (2,017)* %
Strongly agree	15	28
Tend to agree	35	40
Neither agree nor disagree	25	18
Tend to disagree	12	7
Strongly disagree	9	5
Don't know	3	3
Agree	50	67
Disagree	21	12

INTRODUCTION: *For the next question, imagine you have been asked to give guidance on whether access to people's health data should be permitted or not in these situations. You can give a score from 1-5 where 1 is completely unacceptable and 5 is completely acceptable.*

SPLIT SAMPLE A

Q5A. ***A*** **drug company wants to run further tests on an approved drug to see whether the drug has had any unwanted side effects in the long term. The company requests information from a central government collection of medical records, which have had names, addresses and other personal information removed. The company covers the costs of doing this analysis.**

Base:	*(997)*
	%
1. Completely unacceptable	10
2	9
3	24
4	23
5. Completely acceptable	30
Don't know	2
Mean score	3.55

SPLIT SAMPLE B

Q5B. **A public health regulator wants to run further tests on an approved drug to see whether the drug has had any unwanted side effects in the long term. The regulator requests information from a central government collection of medical records, which have had names, addresses and other personal information removed. The regulator covers the costs of doing this analysis.**

Base:	*(1,020)*
	%
1. Completely unacceptable	7
2	7
3	30
4	22
5. Completely acceptable	32
Don't know	2
Mean score	3.65

SPLIT SAMPLE A

Q6A. **To what extent, if at all, would you support insurance companies using health data collected in the NHS to further develop their health insurance prices? This would be done using data with personal information removed, such as name and address. This would mean health insurance prices can be tailored to reflect the risk of ill-health for people living in different areas.**

Base:	*(1,025)*
	%
Strongly support	5
Tend to support	21

Neither support nor oppose	27
Tend to oppose	21
Strongly oppose	24
Don't know	3
Support	26
Oppose	44

SPLIT SAMPLE B

Q6B. **To what extent, if at all, would you support companies using health data collected in the NHS for marketing purposes, for example to target health products at different groups of people? This would be done using data with personal information removed, such as name and address. This could mean, for example, products such as low-fat margarine being promoted in areas where there is a higher risk of heart disease.**

Base:	*(992)*
	%
Strongly support	10
Tend to support	28
Neither support nor oppose	26
Tend to oppose	18
Strongly oppose	18
Don't know	1
Support	37
Oppose	36

SPLIT SAMPLE A

Q7A. **Research using health data can have many benefits, such as developing treatments for diseases or making health services more efficient. However, sometimes this research needs to be done by a commercial organisation who may make a profit out of the research. The information given to the organisation would not include names, addresses or any contact details.**

Which of the following statements comes closest to your view of health data being shared with commercial organisations?

A. I would not want commercial organisations to have access to anonymised health data, even if this means the research does not take place

B. The research should be conducted by commercial organisations if there is a possibility of new treatments for diseases being developed.

Base:	*(974)*
	%
Agree much more with A than with B	13
Agree a little more with A than with B	12
Agree equally with both / don't agree with either	14
Agree a little more with B than with A	31
Agree much more with B than with A	31
Agree more with A	25
Agree more with B	61

SPLIT SAMPLE B

Q7B. **Some commercial data analysis companies are interested in looking at a large number of hospital records, which do not contain names and addresses. Using many records allows analysts to see patterns that enable them to provide analysis to the NHS on how to improve services. However, due to the large numbers, it is not possible for the NHS to ask permission from patients each time a request is made to access the health data.**

Which of the following statements comes closest to your view of commercial organisations seeking to access this kind of anonymised health data?

A. I would rather the NHS ask patients' permission to share anonymised data with commercial organisations, even if this means some of this kind of research does not take place

B. I would rather this research happen, even if in some cases the NHS does not ask for permission from patients

Base:	*(1,043)* %
Agree much more with A than with B	36
Agree a little more with A than with B	18
Agree equally with both / don't agree with either	12
Agree a little more with B than with A	17
Agree much more with B than with A	17
Agree more with A	54
Agree more with B	34

Q8a. **Which of the following conditions, if any, would you have in place before a commercial organisation, such as a drug company or medical technology manufacturer, could access NHS health data for research purposes? If you do NOT want commercial organisations to have access to health data please say so.**

Base:	*All respondents (2,017)* %
Strict rules that the data cannot be passed on to third parties	53
All names and personal information removed from the data before access	52
Storage of the data in a secure facility	47
Criminal sanctions/ heavy fines if companies are found to have misused data	47
Clear intent that the research will lead to benefits for wider society	43
Approval from an independent oversight committee of ethics experts and academic researchers	34

Any use of the data for marketing purposes is made illegal	32
Commercial organisations to be limited in the amount of profit they can earn from the research	28
Other (specify)	*
I don't think any of these conditions are necessary	3
I do not want commercial organisations to have access to health data for research under any circumstances	17
Don't know	1

Q8b. **Which of the following views, if any, comes closest to why you do not want commercial organisations to have access to health data under any circumstances?**

Base (all who do not want commercial organisations to have access to health data under any circumstances):	*(356)*
	%
They cannot be trusted to store the data safely	20
I don't agree any profit should be made from people's NHS health data, even if there are health and society-wide benefits from the commercial organisation's activities as well.	18
Commercial organisations cannot be trusted to put the interests of society before profitmaking	16
They might sell the data onto another commercial organisation and you cannot control where it ends up	13
They may try and market products and services to me	8
If commercial organisations access the data, they could manipulate it for their own benefit and this is unfair	8
There might be negative consequences for me or my family in the future (e.g. information about my medical history are leaked)	6
They may re-identify me even though names and personal information might be removed from the data	2
There might be negative consequences for the community (e.g. lead to the NHS being charged more for certain drugs)	2
Even if they misuse the data, they won't be punished	2
Other (specify)	2
Don't know	4
No answer	1

9.3.4 Full question wording for Q2

Q2. **As you may know, the NHS and other health services collect data about individuals, for example GP records and hospital visits. This data is used for the direct care of patients, but it can also be useful for hospitals, commercial organisations, researchers and others. Allowing access to data can bring benefits, such as enabling research to find more effective medical treatments for example. However, some people worry that allowing others access to health data will be a risk to their privacy and security, for example if it is possible to re-identify individuals. Overall, which of the following statements is closest to your view?**

SPLIT SAMPLE A – OPTING OUT
a) "We should share all the data we can because it benefits health services and me – as long as I can opt out if I choose"
b) "We should not share data, even if I have the right to opt out, as the risks to people's privacy and security outweigh the benefits"

SPLIT SAMPLE B – ANONYMITY
a) "We should share all the data we can because it benefits health services and me – as long as the data is anonymised and I can't be identified as an individual"
b) "We should not share anonymised data as the risks to people's privacy and security outweigh the benefits"

SPLIT SAMPLE C – CONTROLLED ACCESS
a) "We should share all the data we can because it benefits health services and me – as long as there are strict controls on who can access the data and how it is used"
b) "We should not share data even with strict controls as the risks to people's privacy and security outweigh the benefits"

SPLIT SAMPLE D – GOVERNANCE/ENFORCEMENT
a) "We should share all the data we can because it benefits health services and me – as long as there are heavy fines and possible prison sentences for anyone caught misusing the data"
b) "We should not share data even with heavy fines and possible prison sentences for anyone caught misusing the data as the risks to people's privacy and security outweigh the benefits"

SPLIT SAMPLE E – CONTROL QUESTION
a) "We should share all the data we can because it benefits health services and me."
b) "We should not share data as the risks to people's privacy and security outweigh the benefits."

Agree much more with a than with b
Agree a little more with a than with b
Agree equally with both / don't agree with either
Agree a little more with b than with a
Agree much more with b than with a
Don't know

9.4 External contributors to the research

To follow is a list of individuals, outside the core project teams at Ipsos MORI and the Wellcome Trust, who have contributed to the design and/or delivery of this research.

External Advisory Group

Informed the design of the tender and selection of proposals; helped shape the key research questions; provided input on the development of the quantitative phase of research, and feedback during the report drafting:

Bill Burns – former CEO Roche Pharmaceuticals
Simon Denegri – Chair of INVOLVE; NIHR Director for Public Participation and Engagement in Research
Ben Goldacre – clinician; writer; campaigner; Senior Clinical Research Fellow, University of Oxford
Vivienne Parry – Science writer and broadcaster; Head of Engagement, Genomics England
Sir Nick Partridge – non-executive Director, Health and Social Care Information Centre
Sam Smith – coordinator for medConfidential
Daniel Start – Dialogue and Engagement Specialist, Sciencewise

Sources for case study material

Charlie Campbell – Association of British Insurers
Rob Frost – GSK
Shahid Hanif – Association of British Pharmaceutical Industries
Kim Kingan – Aridhia
Antonis Kousoulis – Clinical Practice Research Datalink
Zizis Kozlakidis – UCL
Jenny Westaway – Health and Social Care Information Centre
Paul Wicks – PatientsLikeMe

Participant recruitment

Provided specific support to recruit cohort member and rare disease patient groups

Farhana Ali and Nick Meade – Rare Disease UK
Debbie Lawlor, Lynn Molloy, Madeleine Murtargh and Ross Robinson – Avon Longitudinal Study of Parents and Children

Contribution to General Public Workshops

Provided insights, highlighted key issues and entertained participants at deliberative workshops in Wrexham and Sutton Coldfield:

James 'Lloydie' Lloyd – **Improvisational actor**
Liz Peters – Improvisational actor

Observed participants at the Sutton Coldfield deliberative event to provide detailed anthropological insights, which fed into the writing of this report:

Lydia Nicholas – Digital Anthroplologist/Senior Researcher, Collective Intelligence Team - Nesta

Additional input

Provided advice on case studies or overall direction of research in planning phases:

Jessica Bland – Nesta
Karen Folkes – Department for Business, Innovation and Skills
Sarah Garner – NICE
Nathan Lea – Farr Institute, London
Carol Lyon – PHG Foundation
Anna Middleton – Wellcome Trust Sanger Institute
Malcolm Oswald – University of Manchester
Tim Sprosen – Clinical Trial Study Unit, University of Oxford
Mark Taylor and Victoria Chico – University of Sheffield

Sarah Castell
Research Director
Ipsos MORI Social Research Institute
Sarah.castell@ipsos.com

Harry Evans
Research Executive
Ipsos MORI Social Research Institute
Harry.evans@ipsos.com

For more information

Ipsos MORI
3 Thomas More Square
London E1W 1YW

t: +44 (0)20 7347 3000
f: +44 (0)20 7347 3800

www.ipsos-mori.com
www.twitter.com/IpsosMORI

About Ipsos MORI's Social Research Institute