Sharing health data - lessons for the UK from the European Health Data Space

Rosie Richards[1], Maili Raven-Adams[2], James Maddocks[1], Sarion Bowers[2]

1. NHS Confederation, 2nd Floor, 18 Smith Square, Westminster, London SW1P 3HZ
2. Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, CB10 1SA

Abstract

The European Health Data Space will enable the sharing of European citizens' health data. 25 European countries, including the UK, worked together in the "Towards a European Health Data Space" (TEHDAS) Joint Action, funded by the European Commission, to develop the initial recommendations for realising this complex and transformative infrastructure. Understanding citizens views and concerns around the sharing of their health data is essential to ensuring the success of the European Health Data Space. To this end, a consultation targeting citizens in Belgium, France and the UK was carried out. In this paper we analyse the responses from citizens identifying as from or in the UK, and discuss how the findings from the consultation might relate to ongoing UK endeavours to share health data.

Introduction

The "Towards a European Health Data Space" (TEHDAS) Joint Action is a collaboration of 25 European Countries, including the UK, to establish principles for the creation of a European Health Data Space (EHDS), under the umbrella of the European Commission's Health Programme 2020 and Digital Strategy 2020.

The EHDS is intended to "unleash the full potential of health data"[1] by enabling sharing of European citizens' and associated third country citizens' health data through a framework and infrastructure of common standards, practices; with a unified governance framework aiming to support individuals to have better control of their health data and promote the use of health data in research and innovation.

Recognising the importance of understanding citizens' views on the use of their health data, Work Package 8 of TEHDAS, titled *iCitizens*,[3] sought to improve understanding of citizens' opinions on the reuse of their data and identify ways to engage and inform people on the reuse of their data. Work Package 8 was led by the French Health Data Hub (France) and the National Healthcare Service Center (Hungary), with the NHS Confederation representing the UK and the Wellcome Sanger Institute supporting.

Overall, the consultation revealed that citizens feel intrinsically tied to their own data and want to be respected as a partner in decision-making around the use and sharing of their data.[2] The inclusion of individuals in the governance of their own data could be a valuable way to ensure a working relationship that would foster trust and promote the use of the EHDS.

Policymakers delivering the EHDS and working on UK health data are facing and addressing many of the same issues. For the EHDS to be successful it will need to develop and adopt common standards for data and infrastructure, and the UK can learn from the work of the Joint Action and the development of the EHDS.

From the UK perspective, there are key elements within the consultation responses that have implications for the way health data is used and shared within the UK's health services. The NHS has a vast pool of data[3] and data use for healthcare and health research in the UK is only growing.[4] Partnerships are often formed among a range of stakeholders, involving academic researchers, life science organisations and commercial players, to provide opportunities to advance the our understanding of health and improve diagnosis, treatment and prevention. These advancements provide benefits for society. However, it is increasingly clear that these benefits cannot be fully realised without the support of the citizens whose data is being used and shared.

Participants' attitudes are likely to have been influenced by well-publicised failures of data sharing, such as the 'care.data' scheme which looked to centralise primary healthcare records in England. Concerns around confidentiality, transparency, a lack of stakeholder engagement and the inability to opt-out, resulted in the NHS abandoning their proposal and appears to be impacting the roll out of the General Practice Data for Planning and Research (GPDPR) scheme.[5] Previous studies have shown that the UK public are wary of NHS data handling competence,[6] and the possibility of the private sector having access to their health data makes them less trusting in data sharing possibilities.[7]

Numerous proposed 'solutions' exist as ways to support individual involvement with decision making over the use of their data, such as dynamic frameworks and secure data environments (SDEs).[8,9]

Here, we present an analysis of responses from UK participants in the Healthy Data Consultation; building upon existing work on UK attitudes. From this analysis we provide recommendations on the use of citizens' health data, as the Government continues to drive ways to utilise health data.[8,10]

Results and Discussion

The Our Healthy Data,[6] consultation used a bespoke platform which allowed participants to leave comments or views in free text boxes in response to a variety of questions, or alternatively visitors to the website could choose to be guided through a "quiz" which presented participants with statements about sharing health data which they could score their level of agreement with and at the end they were presented with a "data sharing personality" based on their score. The platform was open for responses between December 2021 and May 2022.

Although targeted to the citizens of Belgium, France and the UK, the consultation platform was open to anyone. The consultation asked participants to provide demographic information and participants identifying as being from or in the UK were included in the analysis below. Nothing on the consultation was mandatory so participants were not obliged to answer all the questions, or complete the quiz, nor were they required to provide demographic information. Answers already submitted were kept.

In total, individuals identifying themselves as being in or from the UK left 1,041 free text comments and 255 completed the quiz, where they were also able to leave comments.

The quiz presented participants with 9 statements which they could score from “strongly agree” to “strongly disagree”:

A) Data should be reused to support public health research, such as improving our understanding of prevention and treatment of diseases.

B) Data should also be reused for health research such as developing medicines, including where private companies and commercial purposes are involved.

C) Data should only be reused to provide individual patient care and ensure the essential functioning of our health care system.

D) People other than you can decide if the research team can use your information (e.g. a board of experts, an ethical committee).

E) I want to be informed about how the research project uses my health data (e.g. by who, for which purposes, for what benefits?).

F) The research team have discovered that people undergoing the same treatment have experienced serious side effects. Would you like the research team contact you with this information?

G) If the research team wants to reuse my health data, the research should only benefit individuals whose data has been reused, i.e. only those living in my city.

H) Sharing de-identified health data should be mandatory for research programmes that support the common good.

I) People who do not agree to provide access to their de-identified health data should pay more for their medical care.

In addition, participants were asked the following questions which they were able to respond with free text:

a) What should your health data be reused for?
b) Under which conditions can your health data be reused? Why?

c) What should be the role of citizens in reuse of their health data for purposes beyond individual care?

Participants mostly agreed or strongly agreed with their health data being used for improving health and healthcare, where there was societal benefit (Figure 1a, b, c). Notably, while there was a strong majority who supported sharing of data for research purposes (89%) (Figure 1a), there was noticeably less support for the involvement of commercial and private companies (67%) in research (Figure 1b). However, despite the drop in support the majority still supported sharing of data for research even where commercial interests were involved, suggesting that the respondents placed a particular importance on the purpose for which the data are being reused.

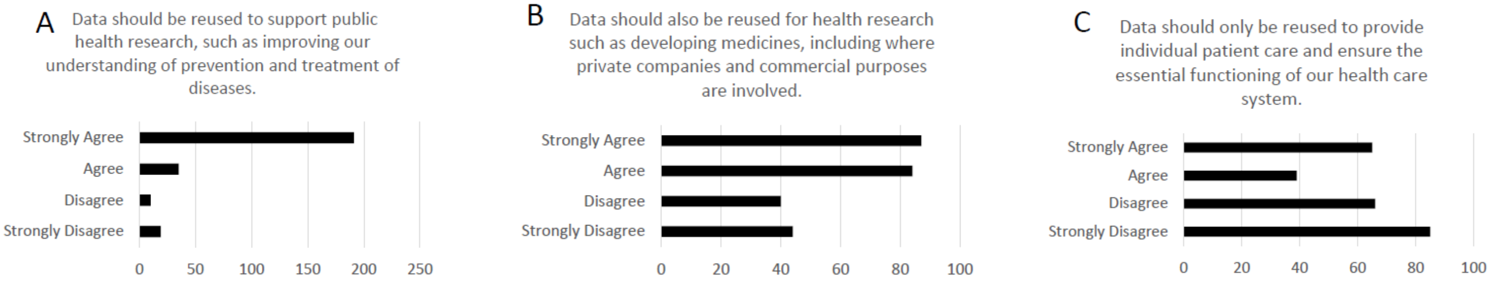


Figure 1a,b,c. Citizen responses to questions posed in the Healthy Data Consultation. Participants were asked to score their agreement with the statements above each graph on a Likert scale of strongly agree to strongly disagree. The X axis shows the number of responses.

Interestingly, despite the strong support for sharing of data for research purposes, the respondents were split when asked if data should only be used for individual patient care or functioning of the health system, with 59% disagreeing and 41% agreeing (Figure 1c). It's not possible to fully explain this discrepancy from the data available, free text responses to the question "what should your health data be reused for?" suggest some citizens view health research carried out in the public sector as a function of the health system. This would be supported by the findings for the recent Campaign for Science and Engineering report on public attitudes to R&D[11] indicating there is a poor understanding of what research is and the purpose it serves in the UK. A varied view on research and what it is may explain the split response from the public.

In parallel to this, there were numerous responses from participants stating their objection to commercial organisations accessing publicly held health data and concerns about companies profiting from the use of citizens' health data. Comments and sentiments such as *"Health data should be reused solely for the purpose of further public research into health and disease. It should not be used for commercial gain"* [anonymous UK participant] were repeated by a number of respondents. Notably, the majority of respondents who mentioned the words "commercial", "company" or "profit" did so in negative terms. The concept of making profit from the use of citizens' data was particularly negatively viewed, but some respondents recognised the need for commercial organisations to be involved in health research but stressed the need for safeguards and benefit to the NHS and citizens.

The respondents were also split on the question of who should decide who can use individuals' health data (Figure 1d). While the majority did agree that people such as an expert panel or a research ethics committee can make those decisions, a significant minority (35%) disagreed or strongly disagreed that people other than them should be able to make that decision. This is a concerning finding given that the decision to share health data for secondary use is routinely decided by ethics committees and data access committees. This finding may reflect the UK’s history with events like care.data and a long-standing lack of trust around the sharing of citizens’ health data. This presents a dilemma for policy makers seeking to implement programmes such as the General Practice Data for Planning and Research (GPDPR)[5].

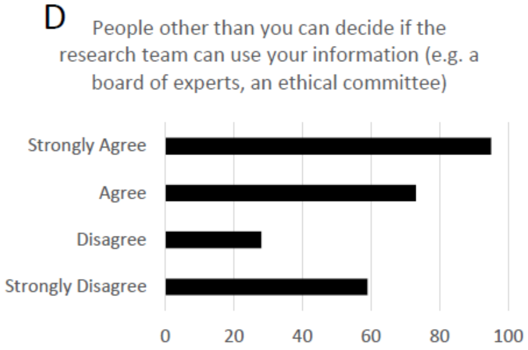


Figure 1d. Citizen responses to questions posed in the Healthy Data Consultation. Participants were asked to score their agreement with the statements above each graph on a Likert scale of strongly agree to strongly disagree. The X axis shows the number of responses.

The questions relating to information being returned to patients from researchers (Figure 1e) and on pertinent findings relating to individuals’ healthcare (Figure 1f), prompted strongly affirmative responses from respondents with the majority wanting to be informed. This diverges from current practice. While some research projects do use dynamic consent and employ other processes whereby they return to participants to ask them for consent to use their data in further research projects or otherwise inform them of additional research, these studies are the exception rather than the norm, and this is not standard with NHS data. There are studies like the ZOE study,[12] examining nutrition and health, which are built around ongoing engagement and input from study participants who engage through the ZOE app and receive personalised feedback. Dynamic consents and returning information to participants/patients are not without ethical challenges, however.[13,14] Incidental findings where the significance is uncertain present a particular challenge to clinician and researchers, and implementing processes for both returning findings and dynamic consent models can be significantly resource intensive, presenting an ethical challenge in itself about use of limited resources. However, the desire of the participants to receive more information and feedback about how their own data is being used is clear, and would seem to be an avenue by which trust could be built between those using data and those providing data.

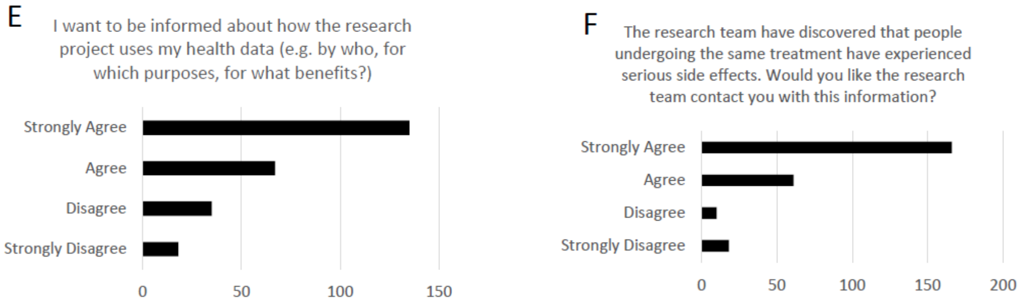


Figure 1e,f. Citizen responses to questions posed in the Healthy Data Consultation. Participants were asked to score their agreement with the statements above each graph on a Likert scale of strongly agree to strongly disagree. The X axis shows the number of responses.

While participants did want to receive information about how their own data were being used, and potentially directly benefit from that use through the feedback of pertinent medical findings, they also expressed a strong sense of altruism that extended to fair and equitable benefits for everyone regardless of the choices of others. There was strong disagreement (87%) (Figure 1g) that only people who chose to share data should benefit from the research resulting from reuse of their data, and that those who chose not to allow access to their de-identified data should have to pay more for their medical care (88%) (Figure 1i). The strong disagreement with these two statements shows a strong sense of altruism amongst participants, and suggests that while the public may have quite a complex and nuanced view on how and when their data should be shared and reused, they strongly believe in societal benefit from that use. Ensuring that research, both in public and private settings, provides a return of benefit to the public clearly should be a priority for policymakers seeking to build support for sharing of health data. In addition, policymakers should seek to ensure policies don't unintentionally disadvantage those who do not engage with research or the reuse of their data.

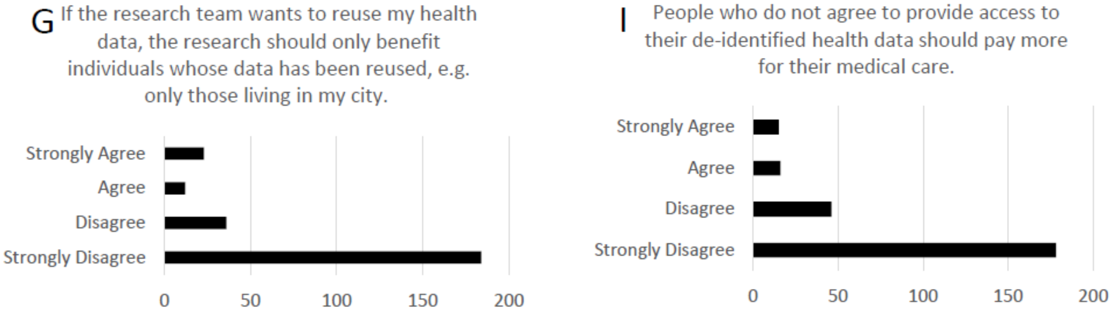


Figure 1g,i. Citizen responses to questions posed in the Healthy Data Consultation. Participants were asked to score their agreement with the statements above each graph on a Likert scale of strongly agree to strongly disagree. The X axis shows the number of responses.

A number of themes emerge from the data gathered during this consultation that are directly applicable to the UK data landscape and new initiatives underway. UK citizens demonstrate altruism alongside a nuanced understanding of what data sharing is and why it is necessary. However, while they recognise the necessary role of commercial organisations, they express scepticism that sufficient benefits are returned to the public from these companies. In addition, while citizens have this nuanced understanding of data sharing, there is confusion around details, with complex terminology and inconsistent definitions obscuring citizens' understanding of what is happening with their data.

There is obviously a need to take citizens' views and preferences better into account when developing new policies, legislation and infrastructures for sharing of health data. Everything from communication to technical processes should be adapted to accommodate the needs and preferences expressed by citizens. This would also include ensuring complex or unfamiliar terminology such as "federation" or "anonymisation" is explained in lay language, a view

Adapting communication to citizens' needs is especially important as many citizens in the free text responses identified concepts relating to federation, anonymisation, de-identification, and benefit sharing as safeguards that would increase their support of the reuse of their data and their trust in those making decisions around the use of their data. Anonymisation was particularly highlighted by participants, showing that it is seen as an essential safeguard by citizens. However, it was also clear that individuals did not always understand the limitations created by anonymisation and there is a tension between the benefit citizens wish to see and anonymisation. Accessible explanations of what is meant by anonymisation, pseudonymisation and de-identification amongst others, and what limitations, risks and benefits are, are necessary.

Citizens placed a particular emphasis on benefit being returned to the public often via the improvement of the NHS and its services, with comments such as "*To help benefit health outcomes for other people as well as myself. My data shouldn't be reused for profit but should be to make better health research and treatments. I expect this to improve the national health service and treatments that come out of research*" [anonymous UK respondent]. Clear communication and evidence of benefit of data reuse is clearly required, especially around the role of commercial organisations in research and development of new treatments. At the same time, it should be clear to the public that decisions to share data are based on considerations of benefit to the public, rather than the potential for profit.

Interestingly, although UK policymakers working on health data have given significant thought to opt-ins and opt-outs, there were only 7 free text comments mentioning opt-out. However, all 7 comments emphasised the importance of having the choice to opt-out. One highlighted a desire to be able to choose what to opt-out of rather than a blanket opt-out, and the remaining comments painted a picture of confusion about what the opt-out means and how it works, with scepticism about compliance. More research would be needed to confirm this result, but it aligns with work done by Understanding Patient Data with the National Data Guardian on public benefit showing that the public wants control over the use of their data.[16] Given the importance placed on being able to opt-out by the public, and the roll out of the NHS smartphone application, "app", which contains the "opt-out"[15], it is imperative that there is clarity for the public on how they can opt-out of their data sharing being used, and what opting out means.

While citizens expressed a clear view of wanting to be informed and involved in the sharing of their data, it was also clear that there was a tension between some of their views. For example, the benefits they wished to see delivered would be compromised by some of the restrictions or processes they emphasised the importance of. The clearest example of this was the conflict between sharing only fully anonymised data and maximising the benefit that can be derived from that data. Similarly, while respondents were clear on the importance of citizens' being involved in decision making around the use of their health data, they also showed a lack of knowledge around the processes and decision making currently in place. For example, one respondent stated "*Although Data Access and Ethical Committees play an essential part in recommending that health data should be shared, all decision-making processes related to re-using health data should involve citizens*" [anonymous UK respondent], this is despite the fact the NHS Research Ethics Committees (RECs) all have lay members of the public on them, as do many other ethical committees. Whether these lay members are truly lay and how best to recruit truly lay members are questions that may be worth further scrutiny. It's clear that it is important to develop strategies that engage the public in all aspects of the use of their data. That could be through engaging the public in co-decision making or by a more unidirectional information sharing approach to explain processes or decision making.

Summary

The findings of this consultation show a public who are able to engage with complex ideas about sharing of their health data and express to policymakers nuanced concepts around the use of their health data. However, there is a mixed understanding of terminologies and the processes that already exist in the UK. The UK public are divided on the involvement of commercial organisations, with a strong concern about profit-making activities but a more nuanced take on commercial involvement where clear benefit to the NHS and therefore UK citizens would be realised. It is therefore vital that policymakers, the national governments, and researchers engage with the public and build systems that reflect these values in order to rebuild public trust in the health data system in the UK.

Recommendations

Although this consultation was carried out as part of the European Health Data Space, the findings from UK respondents are relevant to the UK data landscape and the UK Government, decision-makers in the NHS and researchers. Adoption of these recommendations could support public trust and engagement with the sharing of their health and personal data.

1) *The decision-making processes for sharing data should take a citizen-centred approach from the outset.* It is clear that citizens want and expect to have their views considered and acted upon, and want to be able to access information and knowledge about the reuse of their data. Ensuring systems and processes are designed to facilitate citizen engagement must be part of the design and development phases of projects to share data.

2) *Equity should be embedded into decisions and design.* The public expressed a clear view on the need for demonstrable benefit in order to support sharing of their health data. Embedding considerations of health equity and health inequalities in developing frameworks for considering benefit would further strengthen public support.
3) *Communication on how data are being used and by whom, needs to be continuous.* Although respondents expressed different views on what they wanted to know about the sharing of their health data, it was clear that they wanted to be informed about how their data were bring used, why and by whom. They also wanted to know the outcomes of that use, with societal benefit being clearly linked to public support for the use of health data in research.
4) *Citizens should be engaged at all stages of their data use, to understand their views and involve them in decision making and process development.* There should be a clear understanding of the purpose of the engagement. It should be recognised that citizens' views may change over time and engagement should be able to meaningfully capture the developing views of citizens.
5) *Citizens' views and involvement should be sought for a wide range of factors.* Citizens expressed views on concepts that touched on topics from federation to ethical approval, often reflecting a nuanced view of these complex issues. Engaging and empowering citizens across the breadth of the data-sharing process, would help build public confidence and support for continued sharing of their data.
6) *There must be transparency around decision-making relating to the use of citizens' health data.* Decisions on the use of health data necessarily must take into consideration many factors beyond citizens' views, and may reflect complex issues such as data localisation versus federation, and impact on trade deals. However, citizens should be informed of these factors and should be aware that their views will not be the sole driver of a decision.
7) *There should be clear definitions of benefit, that recognise equity and a return of investment to the NHS that ensure that all citizens can benefit.* Although there was some recognition for the need for commercial organisations to access citizens' data there was a strong antipathy for profit being made from the use of that data. Developing a clear definition of benefit that must be demonstrated in order to receive access to citizens' data would almost certainly build trust in the process and support from the public.

Acknowledgments

With thanks to Nicola Hamilton, from Understanding Patient Data, for reviewing and providing comments and suggestions. RR and SB extend additional thanks to their Work Package 8 partners, Sciensano and the Health Data Hub, specifically Wannes van Hoof, Louise Mathieu and Chloé Mayeur for their work co-designing the Our Healthy Data consultation and platform.

Note: this report often refers to the NHS. This term is used here as a collective for all the health services in the UK, including Health and Social Care (HSC) Northern Ireland.

References

1. Joint Action Towards the European Health Data Space – TEHDAS. *Tehdas* https://tehdas.eu/.
2. maria.nurmi. TEHDAS consultation: Citizens support the secondary use of health data when it matches their ethical values. *Tehdas* https://tehdas.eu/results/tehdas-consultation-citizens-support-the-secondary-use-of-health-data-when-it-matches-their-ethical-values/ (2023).
3. The foundations of fairness for NHS health data sharing. https://www.adalovelaceinstitute.org/blog/the-foundations-of-fairness-for-nhs-health-data-sharing/.
4. Ford, E. *et al.* Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. *Learn. Health Syst.* **3**, e10191 (2019).
5. General Practice Data for Planning and Research (GPDPR). *NHS Digital* https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research.
6. Hill, E. M., Turner, E. L., Martin, R. M. & Donovan, J. L. “Let’s get the best quality research we can”: public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. *BMC Med. Res. Methodol.* **13**, 72 (2013).
7. Stevenson, F., Lloyd, N., Harrington, L. & Wallace, P. Use of electronic patient records for research: views of patients and staff in general practice. *Fam. Pract.* **30**, 227–232 (2013).
8. Better, broader, safer: using health data for research and analysis. *GOV.UK* https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis.
9. Hewitt Review: an independent review of integrated care systems. *GOV.UK*

https://www.gov.uk/government/publications/the-hewitt-review-an-independent-review-of-integrated-care-systems.

10. Data saves lives: reshaping health and social care with data. *GOV.UK* https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data (2022).

11. Public Attitudes to R&D. *CaSE* https://www.sciencecampaign.org.uk/what-we-do/public-opinion/public-attitudes-to-r-d/.

12. ZOE—Understand how food affects your body. https://joinzoe.com/.

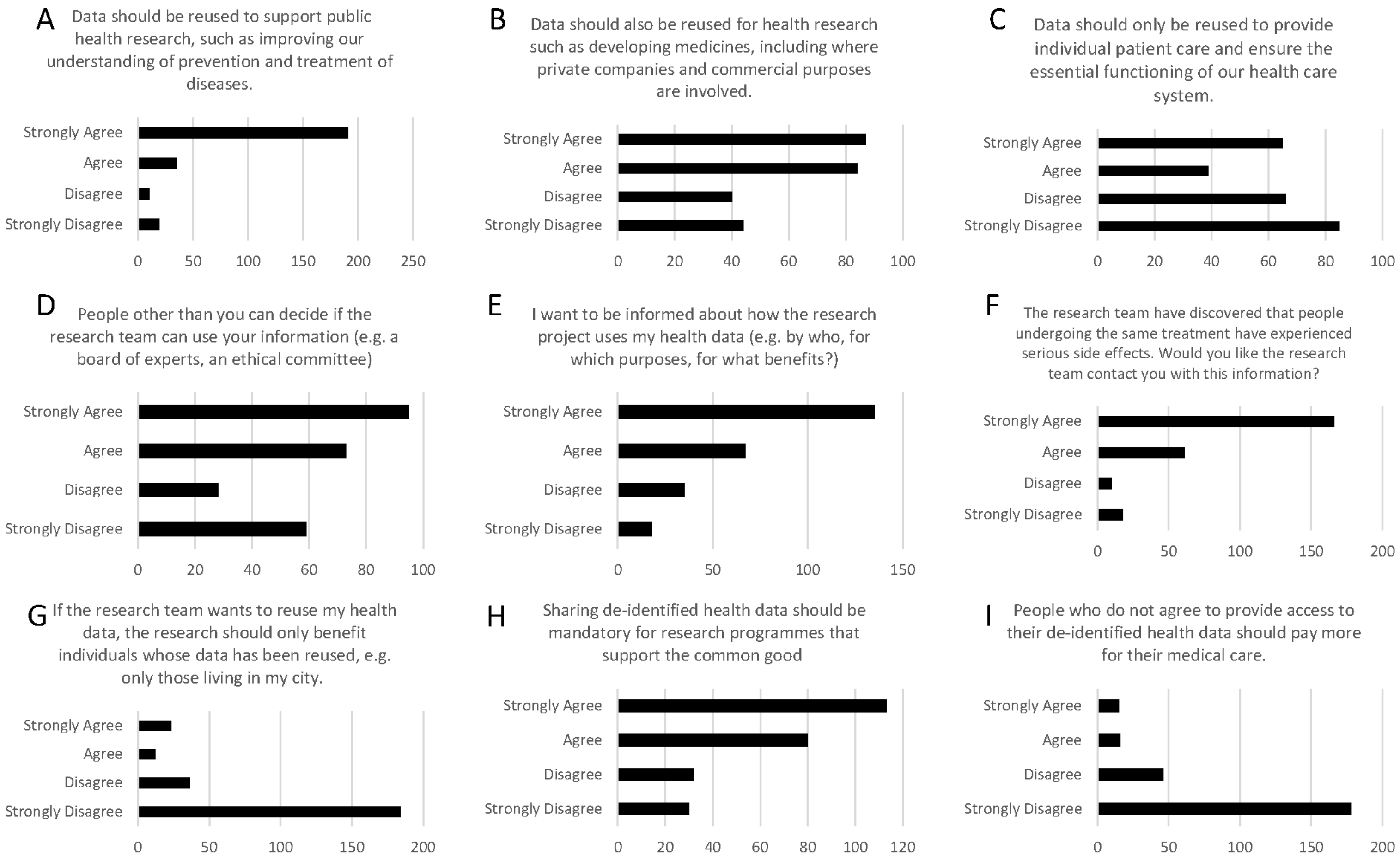
13. Steinsbekk, K. S., Kåre Myskja, B. & Solberg, B. Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? *Eur. J. Hum. Genet.* **21**, 897–902 (2013).

14. Christenhusz, G. M., Devriendt, K. & Dierickx, K. To tell or not to tell? A systematic review of ethical reflections on incidental findings arising in genetics contexts. *Eur. J. Hum. Genet.* **21**, 248–255 (2013).

15. National data opt-out. *NHS Digital* https://digital.nhs.uk/services/national-data-opt-out.

16. What counts as a 'public benefit' for data use? | Understanding patient data. https://understandingpatientdata.org.uk/news/what-counts-public-benefit-data-use.

Appendix A. Graphs


A Data should be reused to support public health research, such as improving our understanding of prevention and treatment of diseases.

B Data should also be reused for health research such as developing medicines, including where private companies and commercial purposes are involved.

C Data should only be reused to provide individual patient care and ensure the essential functioning of our health care system.

D People other than you can decide if the research team can use your information (e.g. a board of experts, an ethical committee)

E I want to be informed about how the research project uses my health data (e.g. by who, for which purposes, for what benefits?)

F The research team have discovered that people undergoing the same treatment have experienced serious side effects. Would you like the research team contact you with this information?

G If the research team wants to reuse my health data, the research should only benefit individuals whose data has been reused, e.g. only those living in my city.

H Sharing de-identified health data should be mandatory for research programmes that support the common good

I People who do not agree to provide access to their de-identified health data should pay more for their medical care.

Figure 1. Citizen responses to questions posed in the Healthy Data Consultation. Participants were asked to score their agreement with the statements above each graph on a Likert scale of strongly agree to strongly disagree. The X axis shows the number of responses.