

Towards
European
Health
Data
Space

Milestone 8.2

Healthy Data, an online citizen consultation about health data reuse – intermediate report

30 June 2022

This project has been co-funded by the European Union's 3rd Health Programme (2014-2020) under Grant Agreement no 101035467.



0 Document info

0.1 Authors

Author	Partner
James Maddocks	NHS Confederation
Louise Mathieu	Sciensano
Rosie Richards	NHS Confederation
Marlies Saelaert	Sciensano
Wannes Van Hoof	Sciensano

0.2 Keywords

Keywords	TEHDAS, Joint Action, Health Data, Health Data Space, Data Space, HP-JA-2020-1
-----------------	--

Accepted in Project Steering Group on 28 June 2022.

Disclaimer

The content of this deliverable represents the views of the author(s) only and is his/her/their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use of its contents.

Copyright Notice

Copyright © 2022 TEHDAS Consortium Partners. All rights reserved. For more information on the project, please see www.tehdas.eu.

Contents

1	Executive summary	3
2	Introduction	5
3	Methods	6
3.1	Inform: Communication campaign.....	6
3.2	Deliberate: Platform	6
3.3	Produce: Analysis	8
4	Results	9
4.1	Participants	9
4.1.1	Healthy Data website	9
4.1.2	Interactive quiz	9
4.1.3	Contributions	10
4.2	Considerations about data and data reuse.....	10
4.2.1	Data is powerful	11
4.2.2	Data is both individual and social	11
4.3	Finding the balance between benefits and risks of health data reuse.....	12
4.3.2	Risks of health data reuse	15
4.4	Engagement	17
4.4.1	Information & Communication	17
4.4.2	Engagement.....	20
4.5	Framework.....	22
4.5.1	A regulatory framework as an overarching safeguard.....	22
4.5.2	Citizen involvement in a framework for health data reuse.....	23
4.5.3	Legal dimension	24
4.5.4	Ethical dimension	26
4.5.5	Societal dimension	26
5	Discussion.....	27
6	Conclusion	29
	References.....	30

List of Figures

Figure 1: Overview of the inductive thematic analysis of contributions to the Healthy Data e-consultation.....	12
--	----

List of Tables

Table 1: Numbers of participation	10
Table 2: Social media campaign results overview	10

1 Executive summary

Introduction

The present report is an overview of participants' opinions gathered through a pilot public e-consultation conducted between December 2021 and May 2022, primarily in France, Belgium and the UK, in order to develop citizen inspired recommendations regarding the reuse of their data and their role in this governance. The Healthy Data online consultation platform aimed to inform citizens about health data reuse, deliberate together with citizens about an ethical, legal and societal framework and produce recommendations for the development of the European Health Data Space.

Methods

To reach a broad audience, a largescale communication campaign was launched, including many partnering organisations who leveraged their networks to bring attention to the initiative. A large body of content was created in different formats, such as text, videos, visuals, an interactive quiz and cases. The content was updated regularly to attract different publics. Every piece of content provided neutral information and was centred around key questions about health data reuse.

Results:

4.244 citizens completed the interactive quiz and 5.932 ideas about health data reuse were shared on the platform. These contributions were thematically analysed to identify the overall narrative and the key values and principles that need to be respected when establishing the European Health Data Space.

Considerations about data and data reuse:

The contributors considered data to be powerful: It is going to exist somewhere with the perpetual potential to be used to do something which impacts the person behind the data. This power can be harnessed for good, but also cause harm, so it is important to be conscious of this power and to respect the rights of the people behind the data. The citizens identified fundamental individual and social aspects in health data. On the one hand, citizens wrote about 'my data', 'I own my data', 'I should be in control'. On the other hand, they would talk about reuse systems, platforms, scientific research, databases, commercial use, ... which can all be defined as fundamentally social in the sense that it is intrinsic to these practices that health data of many individuals is combined.

Finding a balance between benefits and risks of health data reuse:

The purpose behind reuse was a key question for participants. They generally supported health data reuse to support the common good (they mention: improve public health, healthcare, research and policy). Conversely, many contributors were against reuse for commercial purposes. Commercial aspects of health data reuse are often referred to within contributions. If some participants argued that involved private actors can be beneficial in multiple ways, a strong majority shares their concerns regarding their involvement within health data reuse. Participants referred to the need to protect their privacy and to acknowledge the technological limits of data security. In this regard, anonymization was the most mentioned concept within the consultation, but it divided participants on whether data should be anonymized or not. Overall, citizens also believed that the chosen safeguards should depend on the level data has been anonymised. Finally, some participants believed a certain form of citizens' control was legitimate and should be allowed to minimize risks from health data reuse.

Information, Communication and Engagement:

Citizens expressed a broad range of preferences regarding the information they wish to receive, the methods used to communicate with them and the opportunities that are provided with for engagement. Views range from a preference for none of the above either because of an objection towards the reuse of data or their perceived ability to participate in decision making, through to complex processes that would allow for citizens to know in real time what data is being used, who it is being used by, how they are using it and what their intentions are. At this higher level of complexity and involvement, they would also have an ability to make decisions about each and every potential of reuse of data.

In reality though, results suggest that citizens views sit somewhere between these two extremes and allow for citizens to receive information and communication that is personalised to them, and engagement mechanisms that give a place to individual citizens, their representatives, and professionals in such a way so as to allow reuse that supports the public interest, promotes the safe and ethical reuse of data, and is conducted in a transparent and inherently trust building manner.

Framework:

A solid framework for health data reuse could be an overarching safeguard against reuse-associated risks, yet only if citizens trusted the authorities behind it. In case of a lack of trust, however, citizens usually did not trust health data reuse, nor its regulatory framework. A framework should also create transparency and allow for citizen involvement in the governance of health data reuse, if desired. Trusting the authorities behind the framework could decrease the wish to be involved yet being informed about reuse usually remained a necessary condition for public trust. Conversely, a certain distrust regarding the current framework could instigate the wish for a more active involvement of different stakeholders, including citizens.

Participants also expressed specific ideas about the legal, ethical and societal dimensions of a suitable framework. A binding legislation for health data reuse, possibly starting from existing regulations such as the GDPR, should delineate authorised data users and reuse purposes. It should also clarify accountability claims and foresee in preventive and sanctioning measures regarding data abuse. Besides, ethical standards should define an overarching code of conduct. Confidentiality, equality, and individual freedom were mentioned as values of utmost importance. Finally, the framework should respect societal values that acknowledge the collective nature of health data reuse. Solidarity and avoiding discrimination were central concerns within the societal dimension of health data reuse.

Discussion:

Citizens' attitudes towards health data reuse can be described as conditional beneficence, as they perceive it as beneficial at a societal level but also as potentially dangerous.

Several aspects seem to have a major impact on their preferences. First, commercial aspects have a major impact on citizens' perceptions of health data reuse and the framework that should be put in place. Then, whether data is anonymized or not seems to have a fundamental influence over citizen support for a general framework or other safeguards. As a basic rule, health data should be reused for goals that are in line with citizens' values. The overall framework for health data reuse should pursue the common good and include solidarity-based values, while implementing appropriate safeguards.

This conditional beneficence is based on both trust and the fear that health data might be reused against citizens, since complete control is not possible. Trust is of utmost importance and should be earned and fostered by respecting citizens' values. One of these central values is co-creation, in which citizens need to be treated as partners and the framework for health data reuse needs to be co-constructed by them and other stakeholders.

Conclusion:

It is up to stakeholders to develop a framework based on this conditional beneficence, ideally in an open dialogue with citizens. This report serves as a starting point for discussions to translate citizens' contributions about what the EHDS should look like.

2 Introduction

Responding to the European Council's call for the creation of a European Health Data Space (EHDS), the Joint Action "Towards a European Health Data Space" (TEHDAS) was launched on February 1, 2021. Co-financed by the European Commission in partnership with 25 countries, this initiative is part of the third Health Programme of the European Union. Its aim is to develop the future policy, legal and technical framework for the sharing and secondary use of health data in the future EHDS.

The Joint Action has set up the Work Package iCitizen (WP8) to elaborate recommendations on the involvement of patients, citizens, and their representatives in the governance of the EHDS. To achieve this, the Work Package Team launched a public e-consultation on December, 13th 2021 until May 8th 2022 in primarily three pilot countries, i.e., Belgium, France, and the United Kingdom, even though the platform was open to any European contributor, to better understand citizens' perceptions and expectations regarding the reuse of their data and their role in this governance.

A preliminary work was conducted by the Work Package Team to elaborate the content and format of the consultation. This preliminary work comprised:

- An exploratory literature review of 71 scientific articles selected on the Google Scholars' database, to provide a state of knowledge on citizens' perceptions on health data and its secondary use, as well as on the role of citizens in health data governance. This document can be found [here](#).
- 53 stakeholders' interviews with actors from the three pilot countries' health data ecosystems, to understand their experiences, perceptions, needs and expectations regarding health data reuse, as well as their knowledge, expectations and projections of citizens' involvement in health data reuse governance. The summary report can be found [here](#).

Once the consultation closed, the Work Package Team proceeded to a thematic analysis of the contributions gathered, to produce the interim report developed below. The goal of this report is to provide an overview and description of the participants' opinions, and to use it as a starting point for discussion with national and European stakeholders in workshops that will be organized throughout September and October 2022. The aim of these workshops will be to transform the content of the interim report into concrete recommendations for the European Commission, on the involvement of patients, citizens and their representatives in the governance of the EHDS. These recommendations will be submitted in February 2023.

3 Methods

There are many different approaches to engaging citizens on a certain topic. Oftentimes, the topic itself demands a specific method or the involvement of a predefined group of citizens. In this case the topic concerns all citizens on an international level during the COVID-19 pandemic. Therefore, we opted for a large-scale online consultation with a focus on three (pilot) countries and three languages, to be able to maximally use the existing networks of the HDH, Sciensano and NHS Confederation.

The biggest benefit to this method is the capacity to reach a large audience and to offer a low barrier of entry for participation. Online content can easily be distributed, shared, adapted and updated to reach a large and diverse population. We developed an extensive communication strategy to disseminate as much information about the consultation as possible through different already existing networks, as well as through our own networks. The downside to this method is the anonymity in an online environment and the limited potential for a real in-depth dialogue that is characteristic to offline public engagement methods. Therefore, we organized several activities in real life contexts to feed into the online discussion platform as well.

3.1 Inform: Communication campaign

The necessary condition for any deliberation with citizens is for them to be informed about the possibility to participate, the topic and the relevance of the consultation. To attain these goals, partnerships were developed with organisations and publications who already have an audience. An overview of different communication efforts can be found [here](#) (under “dissemination activity report”).

The goal of the consultation was to involve a diverse public. Therefore, the communication strategy was aimed at different groups: patient organisations, schools and publications for teachers, traditional media, social media, newsletters, websites, etc. For every targeted public, the content and messages were adapted to highlight how health data reuse could influence our lives in different ways. To that end, the phrase ‘the second life of our health data’ was coined.

The platform was regularly updated to focus on different examples of health data reuse. With every update, a new communication wave was launched, engaging both organisations and publications who had already communicated about the project and looking for new opportunities.

Throughout the project, citizens were invited to subscribe to our newsletter, which resulted in a dedicated community that was informed about every update on the deliberative platform. In total, 880 citizens subscribed to the newsletter.

In the end, every citizen consultation depends on the goodwill of citizens to participate. Beyond maximizing efforts to communicate to different target groups and tailoring content, it is impossible to guarantee a fully inclusive and representative debate, especially in an online format.

3.2 Deliberate: Platform

To inspire as much deliberation as possible, all the content on the platform was centred around some key questions relating to health data reuse. These questions were identified in the literature review, stakeholder interviews and the ever-growing body of contributions to the platform. The goal of these

questions was not to frame the issues or to limit discussions to these particular questions, but rather to give citizens a starting point and to make the large question of health data reuse more tangible. Citizens were actively invited to go beyond the initial questions and all contributions were analysed as one big narrative to stay as close as possible to the original ideas that citizens wanted to communicate. The three questions were:

- What should your health data be reused for?
- Under which conditions can your health data be reused?
- How would you like to be informed about and involved in the reuse of your health data?

On the platform, citizens could create an account and post their ideas under one of the three questions, or a fourth question that welcomed any other idea related to health data reuse. The participants who created an account would see their contribution published under their first name and the first letter of their last name and provided some basic demographic information to the backend of the platform.

The platform launched with a large body of content. All content was created following a layered approach: an attractive and very easy to understand entry level, followed up by some more complex content and more links for people who wanted to do a deep dive. The initial content consisted of several pages that explained health data reuse, data protection, anonymisation, etc., in addition to videos, visual materials and illustrative cases. For these cases and some of the general content, the materials created by [Understanding Patient Data](#) were used as a baseline to ensure simplified language.

An interactive quiz was developed to allow citizens who might not be comfortable with the publication of their idea about health data reuse in public to still participate. The interactive quiz was designed around three themes (linked to the three main topics on the platform), with three dilemmas per theme, where for every dilemma some arguments for and against were presented to those citizens who asked for more information. After every theme, an open text box was presented to participants, allowing them to optionally contribute an idea to the platform anonymously. The main goal of the interactive quiz was to excite citizens about thinking about health data reuse. To that end, they would get one of eight different profiles related to their decisions about health data reuse that explained what a health data space for everyone would look like following their attitude towards health data reuse.

After the initial launch, three specific content updates were planned to show how health data reuse works in different contexts: cancer treatment, the COVID-19 pandemic and a personal health data vault. Every case was built around the same structure as the platform and ultimately aimed at increasing deliberation about health data reuse in society. Hence, the focus was always on asking questions and indicating core issues and areas of tension (for example, scientific progress and data protection). Every case was presented in a unique way: the cancer case was an [interactive comic book](#), the COVID-19 case was built around an [infographic](#) that could easily be shared and the health data vault case focused on a [project](#) trying to change how health data use and reuse is organized.

Finally, a number of initiatives were taken to discuss directly with different groups of citizens. Several schools and civil society organisations opened their doors to the Healthy Data debate. For each

debate, a specific deliberative method was developed based on a particular topic that was highlighted on the platform. The goal of these debates was always to formulate recommendations for health data reuse, which were added to the body of contributions on the platform.

3.3 Produce: Analysis

Every contribution to the platform was moderated by a member of the Healthy Data team. The rules for moderation can be viewed [here](#). During moderation, team members assigned codes to every contribution to capture their essential meaning. Throughout this process, the team followed the method of inductive thematic analysis as described by Braun & Clarke (2006). The team members were in continuous communication with each other regarding potentially questionable moderation and coding: Weekly meetings were organised to discuss in order to keep coding and moderation consistent. In this first step, codes were very descriptive to stay close to the original meaning of the contributions.

In the second phase of the qualitative analysis, every code was revisited to make sure that each one represented a unique unit with a specific meaning within the overall narrative that developed on the platform. This overall narrative was captured in emerging themes, resulting in a process where initial codes informed initial themes, which in turn influenced the recoding process, as is supposed to be the case in an inductive thematic analysis (Braun & Clarke, 2006).

The overall goal of the inductive thematic analysis was to identify citizens' recommendations for health data reuse. To do this, all contributions – no matter their origin – were considered part of one big dataset and given equal consideration in light of two core research questions:

- o Which basic principles and values about health data reuse are present in citizens' contributions to the debate?
 - o This includes considerations about health data and health data reuse (definitions, preconceptions, evaluations, etc.) as well as references to ethical, legal or societal norms and values concerning health data reuse. In short: what the contributions taught us about citizens.
- o What do citizens want the European Health Data Space to look like?
 - o This includes ideas about the technical, practical, safety, organizational, etc. aspects of health data reuse. In short: how citizens would translate their principles and values into a good EHDS.

It was important to separate these two research questions to allow for equal consideration of every participant's voice: while it is not realistic to expect all citizens to be able to imagine what good practices in health data reuse should look like, every citizen can be considered to be an expert in their own values and principles. Arguably, a well-governed EHDS should look to respect these values and principles above all. Therefore, the first research question was considered to be more important and the analysis for the second question was performed in light of findings about citizens' values and principles.

4 Results

4.1 Participants

Contributions	5.932
Visitors*	24.573
Page views**	38.798
Interactive quiz***	4.244

Table 1: Numbers of participation

*Visitor: single website visits

**Page views: each visitor visits one or several pages on the website, which counts as one single page view

*** The test ran on a different website for which we do not have the data on number of visitors

Phase	Impressions*	Clicks	Click-through Rate (CTR)
1	1.256.181	4.284	0,34%
2	667.655	7.338	1.10%
Total	1.923.836	11.622	0,60%

Table 2: Social media campaign results overview

*Impressions help measure the number of people who have seen a post, even if they did not engage in any way with that post.

4.1.1 Healthy Data website

The website was set up so that several URLs would redirect citizens to the main website (ourhealthydata.eu). For example, in Belgium gezondedata.be and debatdesdonnees.be, in France debatdesdonnees.fr and in the UK ourhealthydata.uk. These pages directed citizens to the main page in their own language.

4.1.2 Interactive quiz

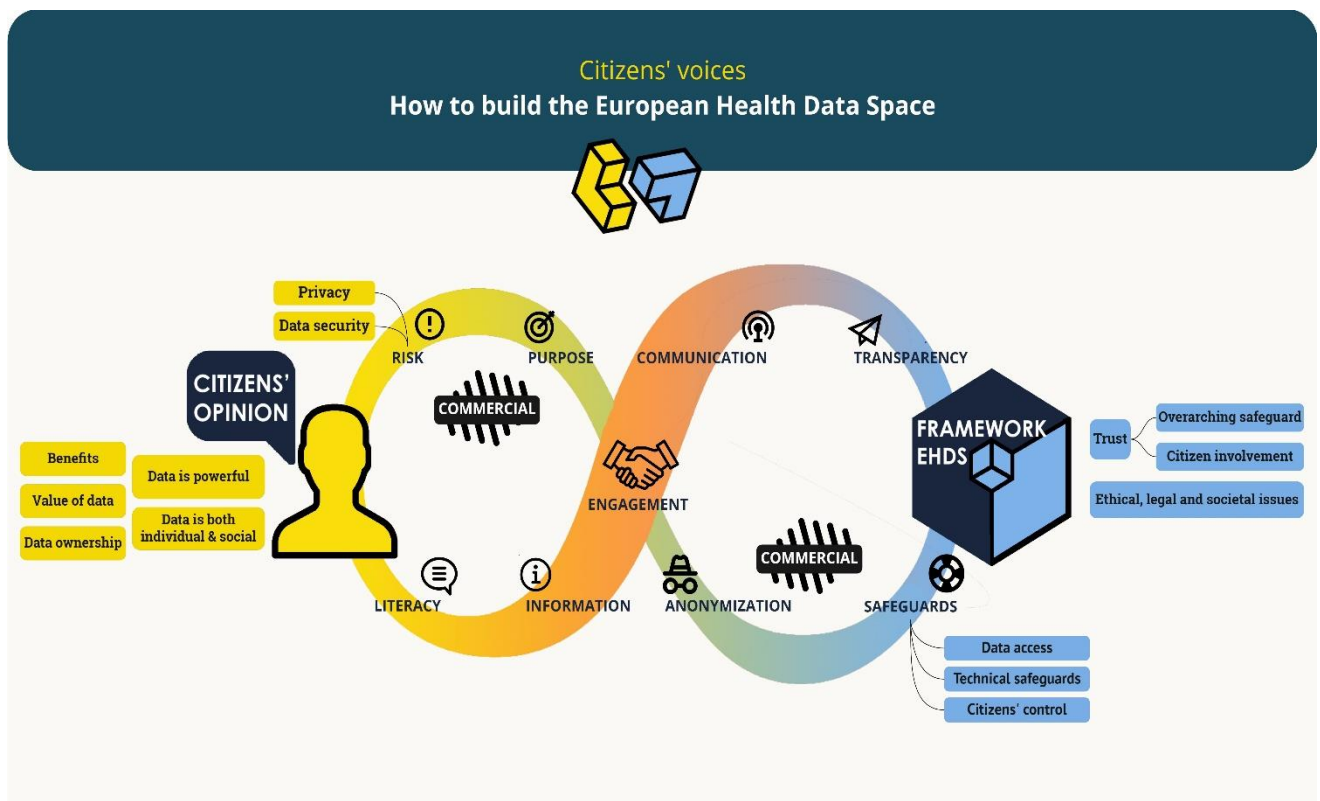
A total of 4239 citizens participated in the interactive quiz; 2410 citizens fully completed the quiz (1176 participants lived in Belgium, 893 in France, 273 in the UK and 168 elsewhere). The goal of the quiz was to inform citizens, to inspire deliberation and to allow citizens to share ideas about health data reuse anonymously. Therefore, the questions were not intended for scientific analysis but an overview of the different responses can be found [here](#).

4.1.3 Contributions

In addition to the citizens who participated through the interactive quiz, 177 citizens created an account directly on the platform. In total, the citizens shared 5932 ideas about health data reuse. After moderation, 5722 ideas were retained. Several contributions were made by groups of citizens (patient organizations, civil society organisations, schools and classes, etc.) and most contributors participated anonymously, so it is impossible to know exactly how many citizens authored one or more ideas.

The thematic analysis of the contributions is captured in this visual overview.

Figure 1: Overview of the inductive thematic analysis of contributions to the Healthy Data e-consultation



4.2 Considerations about data and data reuse

The way citizens conceptualized health data and health data reuse influenced their values and principles in fundamental ways. There were many considerations about data and data reuse, but two underlying themes emerged during the inductive thematic analysis:

- data is powerful
- data is both individual and social

4.2.1 Data is powerful

The starting point for many contributions is the realization that health data reuse is a transaction of power: the data is going to exist somewhere with the perpetual potential to be used to do something which impacts the person behind the data.

Unlocking the power of data is really unlocking the power of people (as the data in question is about people, particularly in health and social care contexts). This highlights the inherent risks of exploitation in using data about people.

Contribution #22708 (UK)

Most citizens argued that this power needs to be harnessed: health data should be stored, reused and even linked and shared as much as possible to explore what we can learn and to reap the benefits. These citizens argued that data should be as open and accessible as possible to maximise its usefulness. Oftentimes, this consideration led citizens to argue for solidarity, where health data sharing and reuse was seen as an individual's civic duty. The citizens who supported these ideas referred to the common good or to specific purposes of health data reuse that aim to address societal needs or improve the lives of everyone (e.g., improving healthcare or scientific advances).

This narrative of data is powerful could also be recognized in the need to keep data secure and reuse under control. The notion of control, whether by citizens themselves, by design or by institutions or committees, was intimately linked to the realization that health data could easily be used for many different purposes by many different actors in impactful ways.

A minority of citizens took this observation of power and need for control to the conclusion that health data reuse should be restricted as much as possible. They argued for hyper strict consent requirements, hermetically sealed vaults or even the abolishing of health data reuse. The key argument behind this position was the assumption that it would be impossible to keep the power of health data under control. No matter what, your personal data can and will be used against you.

Because data is so powerful, it is also valuable, both in an economical and a societal sense. Citizens identified that this value will attract interest from many different parties and result in vast benefits. They argued that these benefits should be distributed equally among the population:

I think we should all contribute to the progress of medicine, but this progress should serve us all and not just the goal of making money.

Contribution #16827 (France)

4.2.2 Data is both individual and social

There was a fundamental tension in the considerations about health data and health data reuse on the consultation platform. On the one hand, citizens wrote about 'my data', 'I own my data', 'I should be in control'. On the other hand, they would talk about reuse systems, platforms, scientific research, databases, commercial use, ... which can all be defined as fundamentally social in the sense that it is intrinsic to these practices that health data of many individuals is combined.

This tension remained when citizens wrote about data ownership: both in an economical sense (buying and selling of health data), a pragmatic sense (use and reuse) and a fundamental sense (who the data belongs to), citizens supported different positions ranging from health data as a most individual good to health data as a societal heritage, sometimes even within one contribution.

The key in this discussion is power (distribution). Medical data belongs to a person, not to an organization (and certainly not to corporatist ones, commonplace in professional sports and the medical community; that is medieval). Government needs to create and facilitate an infrastructure on which individuals may store and manage access to their data (cf. the solid project). This infrastructure offers choice and high-quality services (much like banks in the financial world). It ensures data is never lost or corrupted. It protects against unauthorized access or utilization. [...]

Contribution #17251 (Belgium)

One important distinction that was made by citizens was the difference between health data and the results of health data reuse, where data should be handled with care, but the results should be made public or at least support improved care or other societal benefits.

All forms of data should be used but care must be taken on patient rights and privacy concerns and securing split ownership of intellectual property between commercial entities and hospitals or public authorities.

Contribution #2225 (UK)

The tension between health data as an individual or social concept was often present in the context of risks and benefits of health data reuse. The risks were identified on an individual level (e.g., privacy, discrimination, exploitation) and a societal level (e.g., discrimination, bad quality healthcare or public health, no evidence-based policy, slowed scientific progress). The benefits should be shared equally among everyone in society, but also include individual level benefits where possible (e.g., access to relevant results for one's health). Some talked about rewarding health data sharing and reuse on an individual level, or talked about it like taxes: if you want to profit from the societal benefits, you need to contribute as an individual.

In general, considerations about data were intimately linked to the purpose of health data reuse. All values, principles, ideas and even conceptions about health data reuse changed relative to the context of the purpose. When the purpose was in line with a citizen's values or benefited them directly, all their considerations were impacted. Their data was being used for their ends. When an external purpose was introduced (e.g., profit), a completely different set of rules determined their considerations about data. They were being used as a means to someone else's end.

4.3 Finding the balance between benefits and risks of health data reuse

Many contributions highlighted the benefits that can derive from health data reuse. For some, the benefits actually justify that we need to accept a share of risk when sharing data, or that the current framework is already sufficient to balance these benefits with potential risks. Nonetheless, as stated in the previous part, multiple contributions highlighted citizens perceive risks can emerge from health

data reuse and can be undervalued, or that benefits are low. Another range of ideas dealt directly with privacy, which is sometimes perceived as more important than scientific or commercial benefit.

“I want my health data to be protected in a proportionate way. I know that anonymisation reduces the quality and benefits of the data and I would want this to be balanced carefully. (...)”

Contribution #18885 (UK)

Several ideas mentioned the need to find a balance between benefits of reusing health data and safeguards to prevent potential risks, such as privacy protection. Some contributions referred to balancing mechanisms, such as applying a principle of proportionality or resort to ethical principles in general. A guarantee of a good data management was associated with increasing public trust in that regard within citizens’ contributions, which might explain why some express their wish to be informed about it. 4.3.1 Purposes of health data reuse

Participants’ preferences

One of the main consultation questions on the platform asked citizens what health data should be used for, and since many of them have expressed their preferences, we can have a better view at how they imagined the benefits of reuse being realized. Their preferences on the purpose for the reuse of data were not only mere indications of what could be done, the intention behind the reuse was seen as an important question that should shape the framework behind health data reuse for contributors. For several participants, access to data should depend on the purpose intended by the data user, and hence the goal should be defined beforehand, assessed when requiring access and controlled afterwards to ensure health data is reused for the intended purpose.

“The purpose must be clear and limited, and data must not be used outside it.”

Contribution #24447 (Belgium)

As a consequence, some participants would consider their trust being enhanced if they received the guarantee that health data was reused for its planned goal, indicating a need to receive this information.

On the one hand, improving public health and healthcare seemed to be purposes citizens expected the most from health data reuse, including improving treatments, health prevention and patient safety. Some participants stated that public health should be the primary or even sole goal of health data reuse. For other types of purposes or data users, such as commercial actors, this implied that healthcare systems should end up benefiting from these reuses in some way. Additionally, many citizens mentioned that health data should be reused to improve research and policy. A small minority supported that it should be used for profiling or controlling control people.

On the other hand, some citizens rejected the reuse of health data overall, or the reuse for research, keeping data reuse for essential purposes. Reuse of health data for political purposes, such as for intelligence, surveillance, or to elaborate restrictions, were sometimes also not supported. More consistently, commercial purposes seemed to be the most rejected ones by citizens, including use on social media, marketing, for-profit research, speculation or profiling.

Commercial aspects of health data reuse

Many contributions referred to the commercial aspects of health data reuse, using different references, such as “commercial purposes”, “private organisations”, “private interest”, “industry”, ‘commercial gain”, “private sector” etc. Despite this diversity, participants seemed to refer to two aspects or definitions in the context of data reuse: commercial reuse was either the possibility of generating profits from reusing health data, or the identification of health data as a commodity. The presence of commercial aspects within health data reuse practices gives rise to different points of view.

On the one hand, some contributions referred to the benefits or need of involving private actors in health data reuse, such as in health systems, in treatment production and in innovation in general. Some mentioned the interest of involving these actors in order to compensate a lack of research funding.

On the other hand, the majority of participants referring to commercial aspects described their concern about their involvement in health data reuse.

First, they often associated commercial purposes with a higher risk of abuse. Some referred more precisely to the risk of being re-identified or of data being used against citizens themselves. This potentially conducted several contributors to recommend that health data should not be reused for commercial purposes, or should be limited for commercial purposes, or that private actors should be excluded from data governance, data storage or data processing. Less radically, others commented that safeguards should be increased when the re-user is a commercial entity, or that vulnerable people should be protected with greater care when private companies are involved. Regarding citizens’ engagement, citizen control should be enhanced, according to some participants, when a commercial actor is involved, for instance through the requirement of their consent.

Second, another range of contributions referred to the lack of knowledge about the intentions of private actors, and more broadly to the lack of transparency around the reuse of health data for commercial purposes. This idea was sometimes complemented by the assumption that companies have a financial and/ or political influence that could be exacerbated by reusing health data. Therefore, participants recommended to improve transparency and information towards citizens.

Then, and more specifically linked to the generation of profits from health data reuse, some citizens seemed to put into question the compatibility between health, societal value and public interest with commercial or financial interests. In that regard, various contributors expressed commercial actors should not be the only ones to profit from health data reuse and that benefits should be returned to everyone or should be shared with data subjects, data holders or public services.

“Health data should be used to provide societal benefit, whether that is by public or private companies, but it must be ensured that private companies do provide societal benefit - both through the products they develop and through returning the investment from the public purse and the public donation of data.”

Contribution #17760 (UK)

This could happen through the delivery of free services, the development of affordable treatments, or by imposing financial conditions to access data, such as buying research results or paying for

access, e.g., through a tax. Hence, such a system would require a financial control of benefits generated by data users. Conversely, it is worth mentioning that other contributions, in which commercial actors' involvement was rather supported, indicated that no financial condition should be required to access data.

Finally, when referring to data as a commodity, we identified several concerns about health becoming a financial matter or being monetized. Some contributors advocated against the existence of a health data market or assumed that governments were already selling their data. Two recommended safeguards that could be mentioned as possibly linked to these concerns were the idea that commercial entities should openly publish their results and the idea that private actors should not have the possibility to hold data.

4.3.2 Risks of health data reuse

Privacy and data security

Often, participants strongly linked the requirement of privacy protection to their health and health data, which was often perceived as personal and sensitive information. However, privacy did not mean the same for all participants and some contributors disagreed with this previous statement, as a minority did not see the issue for privacy when their health data are reused. For many contributors, however, health data reuse had a potential to be harmful for them if their privacy is not sufficiently protected. Some of them perceived a violation from the very moment data is reused, when access is provided to identifiable data, or when access is given to entities that are for-profit driven. The main reasons identified were that data could be used against them, or that they could suffer discrimination.

Another common fear among participants was linked to the technological limits of data security, and the potential cyber risks and resulting data breaches. This led some contributors to share concerns about the current data safety, or to assume that full security can never be guaranteed. Technology was also sometimes seen as something to be careful with.

Anonymisation

Besides purposes for health data reuse, data anonymisation was the most dominant topic within participants' contributions. Several contributions show that participants understood the concept as representing a continuum, i.e., that there are different levels of de-identification, such as pseudonymisation. Some of them also thought that anonymisation is only possible until a certain point or hard to achieve, if not impossible.

As for other aspects, participants were divided regarding whether data should be anonymized or not to be reused. For those advocating in favor, re-identification should not be possible, as otherwise this could have a negative impact on data subjects, such as unpleasant news. Anonymisation was also sometimes seen as a security measure per se. Conversely, several contributors asserted that there is no need for anonymisation, and that re-identification should be possible, through a watertight system, Trusted Third Parties, or by medical staff. Contributors justified this view by stating that re-identification could be necessary sometimes and that anonymisation makes data less valuable.

Anonymisation is an important question regarding health data reuse as several participants desired that data access include checking whether data has been anonymized or not. Indeed, it seems that the outcome will have consequences on the framework recommended by citizens. First, anonymisation can impact possible uses of data, as some supported the idea that if data is anonymised, it can therefore be shared and used broadly. Regarding data users, some argued that non-anonymized data can only be accessible to medical staff, or that anonymisation should be the rule when some actors wish to access data, such as private companies. Finally, anonymisation seems to have an important impact on how citizens wished to be engaged within health data reuse governance: if data is anonymized, several did not wish to be either informed or involved, or consent required, and vice versa. Engagement seemed to depend then on anonymisation.

*“When data is not anonymised, the citizen’s consent is required. Otherwise, it is useless.”
Contribution #18494 (France)*

Safeguards recommended

Before considering citizens’ recommendation for safeguards, it is worth mentioning that a minority considered either that there is no need for data protection, or that there are already enough mechanisms in place to protect privacy, such as in the privacy protection regulation. Some also mentioned that they trusted data users in ensuring data security.

Data access

The first set of opinions were referring to actors that can access data or not. It was important for some contributors to put in place a control on the future data user and to filter access to data, as not everyone can access it. Indeed, as mentioned earlier, several participants were reluctant to let private or for-profit organisations access data, in particular insurances, employers, banking organizations, multinationals and the pharmaceutical industry. Politicians were sometimes also mentioned, as well as data users outside national borders or outside the sovereign space in which data has been collected. Finally, public authorities were sometimes excluded from data access. Conversely, a majority of contributors agreed on giving access to actors from the health sector, comprising medical staff, healthcare professionals and organisations, public health researchers and, to a lesser extent, pharmacists and pharmaceutical companies. Universities and medical research experts were also mentioned as possible data re-users. Another set of contributions specified the need that the data users should be independent and trustworthy. Public and governmental entities were also included sometimes. Several contributors mentioned that private companies should not be excluded from data reuse, but some of them added conditions to this access, such as publishing results, developing drugs, not generating profits, or applying for patents.

Technical safeguards

Some participants recommended providing access to only strictly needed data for the purpose intended. Moreover, several of them recommended the use of IT solutions. Regarding privacy protection, they mentioned privacy enhancement technology or secure research environments. More broadly, they seemed to consider that using passwords or codes, accounts, specific platforms, data

encryption or cyber security certificates could enhance data security and privacy protection. A contradiction arises though as to whether it is preferable to use a database or to avoid depending on it and rather focus on protecting the data.

Citizens' control

A final set of safeguards referred to citizens involvement within privacy protection and safeguards to minimize risks from health data reuse. Several citizens considered that since privacy encompasses their health and lives, on the one hand sharing their data should be a choice, and on the other hand it justifies that they should play a role in this framework, as argued in the following contribution:

“The citizen must choose and have free will. No conditions should be imposed on them, they should decide for themselves, their life belongs to them.”

Contribution #21408 (France)

Some self-identified as data custodian, or mentioned that the sole impact on their privacy should entail that data use should be submitted to them, including that requiring consent and providing them information should be mandatory, based on the fact that their privacy is at stake.

Several contributors acknowledged their lack of awareness around safeguards, data security and privacy protection, and they mentioned explicitly their need to be informed about these aspects. This could be explained by the fact they believed that the implementation of security measures and safeguards could help build their trust in health data reuse.

4.4 Engagement

4.4.1 Information & Communication

What do we mean when we say Information, Communication and Engagement?

Information, Communication, and Engagement form the processes through which citizens interact with information, decision making and governance related to health data reuse.

In this context, information describes what citizens are told about reuse and could include why and how data is reused, who it is being reused by, the reasons for reuse, and how citizens' rights will be protected. Communication, in its simplest form, sets out the preferred methods by which this information is shared with citizens; whereas engagement describes the ways in which citizens are given a platform to make decisions about how health data is reused.

Collectively, these three processes form an important strategic function in the development of key data infrastructure and governance processes as they allow for citizens to participate in an iterative dialogue process during the development stage.

What citizens say they want

There were a broad range of preferences given by citizens about what they want from their communication methods. This is a significant finding in itself as we are able to demonstrate the

complexity associated with meeting the desires of some citizens without leaving others feeling somewhere between ignored and over-burdened, which could lead to increased disenfranchisement from data reuse and potentially poor public sentiment.

The communication methods for citizens fell into four broad categories of preference: remote and impersonal, remote and personal, in-person, no communication.

Remote and impersonal

Requiring the least amount of detail and specificity, remote and impersonal methods of communication included social media campaigns, web pages and annual reporting. These channels provide a means for citizens to be informed about the various ways that data is being reused without being informed directly how the specific data they have contributed is being used. These methods, which also included utilising traditional media, advertising and leafleting, appealed to citizens with little interest in data use on a day-to-day basis

Remote and personal

These citizens required more detailed and personalised communication from organisations involved in data reuse. For some, this could include using everyday channels, such as letters and emails; whilst others expressed a need to be able to access information through a dedicated app, or by logging onto a website or portal using a unique ID.

In-person

At the most intense level, some citizens described a need to be informed about any data reuse on a face-to-face basis. For some this would require all information to be provided to them by their GP, whilst others have suggested receiving presentations from members of the research teams who have used their data.

No communication

It should also be noted that a number of citizens described not wanting to receive any information at all about how data is being reused.

These four broad communication categories and the numerous variations within and across them emphasise the personalised approach that needs to be taken with citizens. As we explore our findings in more detail, we will be able to add to our understanding of what this personalisation consists of.

Why are these methods of communication important to citizens?

When describing why the above communication methods were important to them citizens provided insight into the methods themselves, as well as the broader value of communication.

When discussing the value of communication, motivation and transparency were two of the aspects most popularly cited.

Hearing about health data was described as motivating for citizens that are already aware of data sharing, or who have actively chosen to take part in research. These citizens shared that more

information about how data was used would give a sense of involvement that would lead them, and potentially other citizens, to feeling motivated to share their data again in the future.

For those that saw themselves as having a lower level of awareness, the communication methods described were a means for developing a greater understanding of how and why data is used, which citizens think would lead to higher rates of support among the population.

Those who gave a reason for not wanting to be informed about the reuse of their data gave similar reasons to those that did. They commonly said that as long as their data was anonymised/they could not be identified, there wasn't a chance that their data could be commercialised, and that data will be protected by a strong regulatory framework. These were ultimately the same drivers as those who do require information, only that this group (those who need information) required for there to be proof provided to them that these desires were being met, instead of trusting another authority to take this step for them.

Information

When discussing the information that should be provided, citizens' preferences had a similarly broad range to the communication methods that they wanted to see.

This ranged from a need to only know a brief overview of how data had been used over a certain period of time (quarterly, annually), increasing in some instances to a detailed breakdown of who, how and why at a macro scale. At its most detailed, this requirement for information increased all the way up to the ability to have a detailed breakdown of exactly how individual data had been used in research in real time.

The range in detail required for information and communication were strongly linked. Those that described a desire for remote and impersonal communication also didn't require detailed information to be provided, whilst those seeking to know exactly how data has been used were far more likely to want remote and personal, or in-person communication methods.

Why is this information important to citizens?

Throughout the consultation, Citizens described multiple reasons why it was important for them to be informed about health data reuse. One commonly raised reason was so that citizens could see the benefits of their data being used. These benefits could be at the personal level, through the development of new medical techniques and the development of diagnostics and treatments, or it could be at the population level through service improvement by care providers, public health services, or the development of local and national health policy.

Similar to communication, when describing why knowing that the reuse of data has contributed to improvements in health some citizens described it as being a motivational factor for citizens who support data sharing that would ensure their continued support. It can also be a means for gaining the support of those who are currently unengaged in data sharing, or even those who actively oppose sharing.

There were also very personal motivations for receiving information expressed by some citizens and this was so that they can be informed about research that might directly impact upon their health. This could include personalised medical advice being given to them by their healthcare provider, or

it could be receiving more personalised information about the potential side effects of medication that they are prescribed.

Beyond knowing about the positive applications of data reuse, citizens were also motivated to be informed about how health data is reused due to ethical and data security considerations. For example, the reassurance that data is not being used in ways that citizens do not agree with. This can include knowing that there haven't been any breaches of their privacy and that their rights have been respected. For example, they may wish to know about how data is stored and the level of identifiability attached to the data.

“When agreeing to share your data there should be a clear overview of what your data will be used for and who it will be used by as well as how this will benefit e.g., future treatments, studies etc. Also, whether data will be anonymous or can be linked back to you. Data should be kept securely to prevent breaches and you should be informed should there be a breach.”

Contribution #17614 (UK)

There were also concerns raised about the potential of commercial use. As has already been addressed, citizens described both positive and negative outcomes of commercial companies being involved in data reuse. When discussing information, they expressed a need to know that they are not being exploited. This included:

- o Financial benefit of shareholders: a select few should not impede wider benefits to patients, the health service and society as a whole
- o A strong disliking for any potential use of personal health data for targeted advertising

4.4.2 Engagement

How do citizens want to be involved?

When discussing their preferences for engagement, there were a similarly broad range of views presented by citizens about what they want from their engagement methods. Furthermore, it is important to note that information/communication and engagement are not mutually exclusive topics. The methods of communication and information that are preferred would often, if not always, mirror and influence the engagement style that was preferred, and vice versa.

For a proportion of citizens, there was a desire for absolutely no engagement. The rationale behind this varied and included a lack of interest in participating, the perception that engagement would be a burden, and that engagement was irrelevant because they didn't want health data to be reused in the first place. A more commonly held view among citizens was that they felt that engagement was best served through others. This may have been because they felt that they didn't have the required knowledge to participate, or because they trusted professionals, such as their GP and ethics committee members, to make decisions on their behalf.

Those wishing to be actively engaged preferred mechanisms that include one off activities, project specific engagement or continuous citizen participation within the decision-making architecture of institutions.

One-off activities that were mentioned included public outreach campaigns that allow citizens to be consulted on their views towards a specific topic without being specific to any one particular organisation or reuse project. This engagement could also include public dialogue events such as town halls, focus groups, workshops and citizens juries.

Whilst this deliberative mechanism could be seen by citizens as engaging them in data reuse on a general basis, they also saw them as an opportunity to provide citizens with a voice on a more project specific basis. This engagement was seen by citizens as a way to bring them close to, or even directly within the research team and allow them to shape the design of research and its objectives. This would place citizens in a position of decision-making responsibility, for example adding to project approval and other ‘regulatory’ processes.

There was a widely held view that citizens should have a permanent role in the decision-making structures of institutions involved in the reuse of health data. This role could include citizens sitting alongside professional experts on ethics, access, and/or review committees, or it could include separate citizen specific committees being established.

When discussing exactly how citizens should be represented in committees, one preference was for representative organisations to hold the membership on behalf of citizens as they would know more about data reuse than the average citizen and are therefore viewed by some citizens as a more effective representative of their interests. Other citizens used their responses to advocate for a greater role of unattached individuals in these engagement structures, with two popular choices being experts by experience and randomly selected citizens, much like the jury service process.

It should be noted though that whilst several options are presented above, many citizens do not discuss the engagement of individual citizens, their representatives and professionals in exclusive terms. As we have seen already with communication and information, it was common for a layered approach that sees all parties given meaningful role during the engagement process. This position reflects the value many citizens see in bringing together the expertise of many citizens and professional, and is well summarised in the following contribution:

“It would be helpful if citizens can be involved in how health data are used (the focus of research; what is important to them; what should be studied), but they should not be able to prevent research that is considered important and relevant by expert groups or that is considered feasible by ethical review boards.”

Contribution #22284 (Switzerland)

These methods for engagement would give citizens a means for being engaged on a consistent basis, a requirement that is also seen when discussing information and communication. Whilst specific time frames were not always mentioned, it is apparent that public outreach campaigns and deliberative discussions are seen as giving citizens a means to participate in debate from the outset. Meanwhile, it was noted that the more formal structures favoured by citizens offer close to real time continuous engagement on data use.

Why do citizens want to be engaged in this way?

Ethical oversight, adding to research, and building trust were the most common drivers of citizens' desire to be engaged in data reuse.

When discussing an ethical oversight, citizens described a scenario where they were able to guide the design of projects in a direction that they saw as representing an appropriate reuse of health data. This direct overview and ability to influence research for ethical purposes extended into the active delivery and evaluation of research. One example of how this might be applied is in the event of any data breaches. In such instances citizens described engagement as being a mechanism to become aware of such instances and therefore ensure that appropriate steps/sanctions can be taken.

Citizens also saw themselves as having something valuable to add to research beyond supplying data through the provision of input and insight that will make the research itself better as well. This ranged from guiding researchers to topics of most interest to the public, to using the results to shape public policy.

Shared decision making by engaging citizens in the decision-making process was also seen as a method for promoting transparency and building trust in data sharing and reuse.

4.5 Framework

4.5.1 A regulatory framework as an overarching safeguard

People did not only refer to specific conditions and safeguards to manage health data reuse, they also referred to a more global framework that should combine their concerns. This framework was presented as an overarching regulatory system that integrates, aligns and organizes their considerations about health data reuse. Some participants only referred to the need for a framework in general, without clarifying its more specific organization. This general call for rules and regulation suggested a need for an overarching safeguard.

On the one hand, a clear regulatory framework was often considered a proper safeguard, that in itself enhanced public trust in health data reuse.

Good information governance is not in itself a barrier to data use. Rather, it is essential to ensure that patient data is used in a secure and privacy protecting manner. This is fundamental to demonstrate trustworthiness and therefore build and maintain public trust.
Contribution #19462 (UK)

Public trust in health data reuse and its outlining framework seemed to be enhanced when people had witnessed good examples of health policy and healthcare management, for instance regarding health data security. Also trust in the regulatory authorities behind the framework, such as governmental institutions, experts or patient groups, contributed to a general trust in health data reuse.

In line with the importance citizens attribute to the purpose of health data reuse and the parties by whom these purposes are realised, trusting the framework for health data reuse was impacted not only by those authorities who regulate and define the framework, but also by those parties who are,

within this framework, allowed to reuse health data and by the purposes for which these parties are allowed or expected to use the data. Potential data user should be trustworthy, competent in data management, and demonstrating integrity. Moreover, they are expected to use the data as intended, in a valuable way, or in the best way possible to improve healthcare.

On the other hand, a considerable amount of people did not trust the parties that define the regulatory framework, such as the government, public authorities, or the healthcare system, and hence, they did not trust health data reuse in general. Also, when commercial parties would be able to reuse health data, there was a remarkable decrease of trust in health data reuse. Some of the participants based this lack of trust on negative experiences with health data reuse or on bad examples of health data governance or healthcare management. Other participants explicitly went beyond individual experiences and personal statements and, instead, considered their lack of trust in health data reuse a general, society-wide challenge.

4.5.2 Citizen involvement in a framework for health data reuse

Citizens' references to the need for a clear framework for health data reuse did not only indicate a general call for an overarching safeguard. Additionally, these references included a general question for clarity and transparency regarding the framework in which health data reuse takes place. Therefore, it seems that the call for a framework was also a request to increase public awareness about or involvement in health data reuse and the way it is regulated.

A variety of statements on how this awareness should be realised and to which extent citizens should be involved in the establishment of a framework for health data reuse, were expressed.

Leave it up to the experts

A part of the contributors did not want to be actively involved in the development of a framework and rather preferred to leave this up to, for instance, experts or public authorities. Rejecting the idea of being involved in the establishment of a regulatory framework might be explained by participants' trust in authorities and experts to handle health data reuse appropriately.

However, several participants still claimed the need to respect individual choices within this framework, e.g., through consent. Nevertheless, the demand for choice and engagement was mostly limited to own, personal cases.

Trusting the regulatory framework for health data reuse as well as the decision-making authorities might reduce the wish to be informed about health data reuse. Indeed, some contributors expressed that if they trust bodies responsible for health data reuse management, or if there is a trustworthy framework in place, they do not need to be (abundantly) informed about it or they even do not wish to be asked for consent when their health data would be reused.

In general, however, trusting the framework in which health data reuse currently happens and the authorities that define or act within this framework, did not stop participants from expecting transparency and communication regarding health data reuse and its regulatory frameworks. Even though citizens might put their trust in expert authorities, they still demanded information, communication and transparency concerning these authorities' decisions and actions to build and maintain trust.

Hence, whereas for some participants trust seemed to be perceived as a sufficient condition to reduce information and communication, for most participants information and communication remained necessary conditions for trust.

Actively involving citizens

In comparison with the above-mentioned contributions, many participants wanted to be involved more actively in the development of a framework for health data reuse.

This wish for engagement could be related to a certain degree of distrust regarding current frameworks for health data reuse, as it was mentioned by a group of participants. They believed, for instance, that health data reuse frameworks are defined by too homogenous groups of experts, and they deplored a lack of interdisciplinarity. Instead, decision-making authorities should integrate different perspectives and reflect a multidisciplinary vision, such as through the designation of multidisciplinary committees. Regulation should then be the result of cross-disciplinary teams, including, for instance, health professionals but also lawyers and data protection officers. Such a multidisciplinary approach would be most suitable to define the best possible protection mechanisms or ensure a critical approach within health data reuse, as some contributors claimed.

For a considerable group of citizens trust in health data reuse could, however, only be realized if also citizens are actively involved in the development of a regulatory framework. Information and communication on health data reuse, its regulatory frameworks and the involved experts or authorities (even if they are considered capable or trustworthy) were considered not sufficient and more active forms of citizen engagement were required. For these people, the realization of trust by being actively involved in the development of a regulatory framework, is an essential condition for health data reuse. They wished to be engaged in decision-making processes or policy making, for instance by defining criteria for reuse or designing local data governance structures. Citizens also expressed their desire to participate in data use governance, for example through citizen committees, or to provide advisory opinions or formulate a veto on certain reuses.

Therefore, these participants did not merely want to be able to move freely or safely within a valuable framework that regulates health data reuse. They also (partially) wanted to help draw the lines of this framework. They did not only want to decide on personal, individual cases of health data reuse but also be involved on a higher policy level.

Besides these overarching thoughts on a framework for health data reuse as a general safeguard and on the degree to which they want to actively engage in this framework, participants also shared their insights on more specific issues that should be covered by a regulatory framework. They mainly referred to legal, ethical and societal issues that are highly at stake in health data reuse.

4.5.3 Legal dimension

Legislative context

Some people considered regulations and legislations currently in effect as already suitable to adequately manage health data reuse. They specifically referred to privacy, confidentiality and data protection laws and they claimed that, if well interpreted, aligned and applied, this body of laws does not need to be revised and is fit for health data reuse.

The idea was repeatedly mentioned that the current GDPR legislation is a good starting point to regulate health data reuse. It guarantees confidentiality, privacy, data protection, the opportunity of

personal consent, and transparency, so participants said. Besides the GDPR, participants also supported other guidelines that could serve as good examples for a responsible health data reuse, such as wider health regulations. Therefore, these contributors indicated that a legal framework should not appear from nowhere yet align with existing guidelines.

Others, however, expressed the ideas that there is an unfulfilled need for a binding legal basis for health data reuse and that legal conditions and restrictions require further development. Current strategies and legislations were considered too much siloed and lacking harmonization, resulting in overlapping laws that are not in lockstep and, hence, confusion. Instead, there should be a single, consistent authoritative set of streamlined principles everyone should commit to and cross-sector guidance to create clarity about the framework.

Furthermore, several participants insisted on improving international coordination and finding a harmonization of international regulations.

Responsibilities and accountability

A suitable framework for health data reuse should, according to many participants, define legal restrictions concerning the type of actors or authorities that can reuse health data, the purposes they are serving, and the types of health data that can be reused (e.g., data collected through wearables or mobile applications). The framework should also guarantee that specific rights are respected, for instance citizens' right to change their mind (by installing dynamic reuse agreements) or their right to be forgotten. Another legal issue that was frequently mentioned, was the question of accountability. In the development of a framework for health data reuse, it should be clarified who is accountable for proper health data reuse, such as the government, medical professionals or expert data holders.

This idea was closely linked with the many contributions on actors or authorities that are considered trustworthy and should be responsible for providing access to health data, controlling justified access to data, and following up on actual health data reuse. Suggestions varied to a great extent, ranging from (public) health actors or institutions (such as treating physicians, hospitals, medical colleges, or public health authorities), over public authorities (such as the government, public authorities, or parliamentary assemblies) to (scientific) experts in different aspects of health data reuse. Multidisciplinary committee were also suggested, as well as specific data protection, surveillance, or ethics committees.

Some contributors included citizens as part of the body or authority that should control access to health data and follow up on its reuse. However, not many references were made to citizen involvement regarding legal issues and this domain seemed to be mainly the responsibility of professional experts.

Abusive use and sanctions

As health data was perceived both sensitive and powerful, participants considered it inherently vulnerable to abuses. Particular contexts, in which, for instance, private companies, the government or political interests are involved in health data reuse, were considered to even increase this risk of abuse. Therefore, a good framework including adequate legal measures should be in place as a safeguard to counter these risks and to avoid citizens to become the victim of data abuses.

The system must be strictly regulated to avoid abuse. Obviously, it is out of the question that this data fall into the hands of employers or insurance companies, even if the insurance

companies linked to mutual insurance companies already know almost everything about my health!

Contribution #19445 (Belgium)

Citizens made some recommendations in terms of legal procedures to protect data against unwanted or unjustified eyes. Participants suggested that clear protocols and strict regulations regarding data access, as well as ICT security and privacy protection measures should help avoiding abuses. Transparency about and control over data reuse could further complement these protective measures. Authorities that should be responsible for the development of legal mechanisms to prevent health data abuses included, again, professional experts and specific institutions that should manage and supervise health data reuse.

In spite of the need for strong preventive measures, several participants referred to exemplary cases of previous abuses, in which health data were, for instance, accessed by unjustified parties or used for improper purposes. In those cases of abuse, strong and legally founded sanctions should be available, such as financial compensations or the exclusion of any future health data reuse.

4.5.4 Ethical dimension

A suitable framework for health data reuse should not only consider legal issues, it should also include ethical concerns and values. Various participants considered ethics more important than, for instance, commercial benefits and some citizens even considered the ethical dimension the priority of a valuable and sustainable health data reuse. The underlying reason for these concerns seemed to lie, again, in the specific characteristics and more specifically the sensitive nature of health data.

Even though the ethical dimension of a framework for health data reuse may also function as a safeguard, participants referred less to rules, abuses, or sanctions in this context. Instead, they rather referred to concerns, ethical standards, values, etc. Rather than as a safeguard, participants seemed to consider the ethical dimension of the framework as an overarching code of conduct that should be respected in health data reuse. Important values that were repeatedly referred to included confidentiality and equality, as well as respect for persons, individual freedom and patient rights. These values were very frequently translated into ideas regarding privacy and consent.

In line with concerns regarding the development and implementation of legal mechanisms to manage health data reuse, participants also referred to specific expert authorities, such as independent organizations or ethics committees, to develop, install and track ethical values and dimensions regarding health data reuse. However, contributors referred significantly more to citizen involvement and concepts of collaboration within this ethical dimension of the framework and several expressed a wish to be involved in ethics committees or democratic debates.

I would like to make sure that patients are properly represented in ethics committees that decide on the reuse of data.

Contribution #19309 (Belgium)

4.5.5 Societal dimension

Finally, participants referred to collective values that should be protected in health data reuse and to concerns that indicate how society might be affected by health data reuse. These ideas counterbalanced the many references to individual rights and risks and, instead, reflected ideas of

collective involvement and engagement in health data reuse. Solidarity was mentioned as a necessary central value in a framework for health data reuse and the idea was expressed that reuse should be led by collective decisions.

There should be as wide a variance and quantity of samples of data included in the research to give a balanced result. In practice this will likely mean using some parameters to separate the data into cohorts to ensure that any variance in areas does not skew results. No, no-one should have to pay more for their healthcare under any circumstances.

Contribution #22259 (UK)

Another recurring societal concern regarded the risk of (increased) health inequalities that might be brought about by health data reuse. Hence, vulnerable people should be protected against adverse consequences of health data reuse and, more generally, the risk that citizens would be discriminated because of incorrect or abusive health data reuse, should be avoided as much as possible.

5 Discussion

Citizens' health data reuse ecosystem

The inductive thematic analysis shows how it can be possible to construct an ecosystem for large scale health data reuse based on input from citizens. As Figure 1 shows, two main flows of interlinked themes emerged:

- o Flow Green left to right: How do citizens' core values and ideas translate into a framework for a health data space?
- o Flow Green right to left: How can a health data space respect citizens' core values and principles?
- o Flow Orange left to right: How can citizens be involved in a health data space?
- o Flow Orange right to left: How can a health data space foster trust and listen to the voices of citizens and patients?

The links between the themes flow in both ways because ideally citizens' considerations feed the framework for the health data space and in turn, this framework influences citizens' considerations about health data and health data reuse.

Conditional beneficence

The results from the consultation indicate that citizens can feel strongly about data reuse and linkage because they think data is powerful. Therefore, data can be very beneficial at a societal level, but its use can also be perceived as dangerous, which makes it necessary to set up a form of control. Whether citizens see the balance tipped more towards the risks of re-use, or more towards the benefits of reuse, leads to the same conclusion: data reuse is a balancing exercise where one has to try to understand which benefits are expected and which conditions should be fulfilled to avoid

reuse-related risks. Therefore, citizens' attitudes towards health data reuse can be described as conditional beneficence.

First, references to commercial aspects of health data reuse are very common within contributors' opinions. Participants associate commercial aspects with a higher risk of abuse when reusing data, or they fear that health could become a financial matter. They put into question whether commercial purposes can ever be aligned with the common good. Overall, participants apply a different framework whether the data user is public or private: it has a major influence on citizens' perceptions of health data reuse and the safeguards that should be put in place.

The second condition to mention is data anonymization or identifiability, the most recurrent topics across participants' contributions about risk mitigation. Even though anonymization could be considered solely as a safeguard to protect privacy, whether data is anonymized or not seems to have a fundamental influence over citizen support for a general framework or other safeguards. Whether data is anonymized can impact the support for specific uses or data processors, and will also have an influence on how citizens wish to be involved within the overall health data reuse governance. Regarding the latter, there seems to be a correlation between identifiability and involvement, i.e., the more identifiable people are, the more they wish to be involved, such as by receiving information, giving consent or having more control on the reuse of their data.

An equally important condition is a purpose that is in line with citizens' values. Participants mentioned their criteria to access data either through the purpose intended or the user behind the reuse. When comparing contributions related to the former or the latter, opinions are matching most of the time. Therefore, we could consider that when designating a data user, contributors assume that one type of user will always serve one purpose, e.g., a commercial actor only has commercial purposes, a public health researcher only has public health in mind, etc. The key question then becomes which purposes are in line with citizens' values. At the surface level, they mention reuse for improving health, healthcare, scientific research and (health/ healthcare) policy, but they remain divided when it comes to commercial purposes.

The overall framework for health data reuse should facilitate the pursuit of the common good. In that case, citizens can imagine a solidarity-based framework where health data reuse is self-evident, like taxes. This framework would implement appropriate and proportional safeguards that balance the need for evidence-based health care and policy, the aim of scientific progress and the protection of citizens whose data is reused.

Trust

Conditional beneficence is based on trust: all of the conditions are based on the fear that health data might be reused against citizens, but many citizens realize that complete control or full anonymization is not possible. There will always be a potential for abuse. Therefore, they need to be able to trust that their beneficence will not be taken advantage of. Trust is not the baseline, but something that needs to be earned, developed and fostered. Trust comes with strings attached: the need to respect the values and principles held by citizens regarding health data reuse.

The only way trust can be earned is through transparency. This implies clear information through proper communication channels and openness and honesty about health data reuse practices. Most

citizens do not blindly trust health data reuse, but want to be able to trust it because of all the good it can bring to our society. In essence, the best way to foster this trust is to treat citizens as partners.

Citizens as partners

Citizens need to be treated as partners in health data reuse governance. They are the ones who are being described by the data and whose lives will be influenced by the results of health data reuse. On the platform, many different, sometimes even contradictory strategies to involve citizens were suggested. In that regard, the contributions about information and communication mirror those about involvement, ranging from no need to be involved at all to demands for complete individual control. Generally, whatever type of engagement citizens support, the key thing to note here is that the need for data to be used in a way they support is omnipresent, as even when citizens indicate no need to be involved or informed, this means that they are happy with the status quo, are disinterested, do not support health data reuse at all or trust the existing framework. When citizens are treated as partners, the ethical, legal and societal dimensions in a suitable framework for health data reuse can truly be co-constructed by data subjects, data users and data managers.

6 Conclusion

The conditional beneficence of citizens towards health data reuse requires a well thought out framework with adequate attention to its ethical, legal and societal dimensions. It is up to the policy makers, experts and stakeholders to develop this framework, ideally in an open dialogue with citizens. This report serves a starting point for discussions with European policy makers, experts and stakeholders about the European Health Data Space. It is up to them to translate citizens' contributions about what a health data space should look like into a reality.



References

Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101.