



Towards
European
Health
Data
Space

Deliverable 8.1

Qualitative study to assess citizens' perception of sharing health data for secondary use and recommendations on how to engage citizens in the EHDS

31 March 2023

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
Kenann Menager	Health Data Hub
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 February 2023. The European Commission gives final approval to all joint action's deliverables.

Acknowledgments

We acknowledge the work put of the TEHDAS review board and the valued input from Sarion Bowers (Wellcome Sanger Institute) and Linda Abboud (Sciensano).

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 © 2023 TEHDAS Consortium Partners. All rights reserved. For more information on the project, please see www.tehdas.eu

Table of contents

Abbreviations	4
1 Executive Summary	4
2 Introduction	6
3 How to read this report?	8
4 Recommendations	9
4.1 The data relationship	9
1/ Citizens would recommend being able to access information about the secondary use of health data, in an understandable way, allowing them to be more engaged	11
2/ Citizens would recommend having access to their data and know how they are used for secondary purposes. However, they want to choose how and when they are informed about the uses of their data.	14
3/ Citizens would recommend that their values should inform what is beneficial to individuals and what constitutes the common good.	15
4/ Citizens would recommend that decision-making processes rely on a plurality of views and actors to increase their trustworthiness, as for them the latter depends on who is involved in these instances.	17
5/ Citizens would recommend being given the opportunity to be involved in the lifecycle of health data, as they need to be engaged on a continuous basis. Otherwise, their relationship with data custodians and users can deteriorate.	18
4.2 Power balance.....	21
6/ Citizens would recommend being provided with the opportunity for meaningful and active decision-making in the secondary use of health data, as they value the ability to exercise control	22
7/ Citizens would recommend to ensure the protection of individuals' identity, which they perceive as one of the best ways to balance the harms and benefits of the secondary use of health data.....	25
8/ Citizens would recommend that data users' intentions should be transparent and in line with purposes citizens support, as they think some users might share citizens' values more than others	26
9/ Citizens would recommend that accountability could be enhanced through transparent and stronger mechanisms	28
10/ Citizens would recommend to foster good IT solution to protect their data, beyond having a strong legal framework in place.....	29
4.3 A citizen powered framework.....	29
11/ Citizens would recommend that stakeholders respect principles that align with citizens' ethical values.....	30
12/ Citizens would recommend having a dynamic framework which facilitates the secondary use of health data for purposes and benefits that they support, while minimising the potential risks they identify.....	34

4.4 What could this mean for the EHDS?	37
Informing citizens and raising awareness in an educational way, about the secondary use of health data.....	37
Determining and including the value of the common good in the secondary use of health data	39
Fostering a plurality of views within the framework on the secondary use of health data .	39
Treating citizens as equal partners	40
Identifiability of citizens.....	41
Sanctions and safeguards	41
5 Towards a trusted European Health Data Space	42
Annex 1: Participative tools to the Healthy Data Consultation	43

Abbreviations

EHDS: European Health Data Space

HDABs: Health data access bodies

1 Executive Summary

This document is the final report from the TEHDAS iCitizen Work Package (Work Package 8) and presents citizen recommendations on how they want their data to be used and their desired role in the future European Health Data Space (EHDS). The recommendations summarise the key findings of the qualitative analysis of citizens' contributions to the Healthy Data consultation. The aim of the public consultation was to better understand citizens' perceptions and expectations regarding the secondary use of their health data and their role in its governance. These citizen recommendations are also informed by the Healthy Data literature review, expert interviews and national and European stakeholder workshops.

According to the citizens who responded to the Healthy Data consultation, every time their data is used, they feel that a piece of them, their identity and history, is being used. The consultations evidenced that this sentiment also applied in the case of secondary use of the citizens' data. Therefore, anyone using citizens' data is engaged in a relationship with them and citizens feel like they need to be respected as a partner in this relationship. Every element of the data relationship needs to be given proper attention and requires continuous work.

1. Citizens would recommend being able to access information about the secondary use of health data, in an understandable way, allowing them to be more engaged.
2. Citizens would recommend having access to their data and know how they are used for secondary purposes. However, they want to choose how and when they are informed about the uses of their data.
3. Citizens would recommend that their values should inform what is beneficial to individuals and what constitutes the common good.
4. Citizens would recommend that decision-making processes rely on a plurality of views and actors to increase their trustworthiness, as for them the latter depends on who is involved in these instances.
5. Citizens would recommend being given the opportunity to be involved in the lifecycle of health data, as they need to be engaged on a continuous basis. Otherwise, their relationship with data custodians and users can deteriorate.

The second central theme identified in the contributions to the consultation is the perception of data by citizens as being valuable and powerful: data can be used to greatly benefit everyone, but this power could also be used against them. They communicated the need for a good balance between risk mitigation and benefit maximisation to ensure proportionate secondary use of health data for purposes that they support.

6. Citizens would recommend being provided with the opportunity for meaningful and active decision-making in the secondary use of health data, as they value the ability to exercise control
7. Citizens would recommend to ensure the protection of individuals' identity, which they perceive as one of the best ways to balance the harms and benefits of the secondary use of health data.
8. Citizens would recommend that data users' intentions should be transparent and in line with purposes citizens support, as they think some users might share their values more than others

9. Citizens would recommend that accountability could be enhanced through transparent and stronger mechanisms
10. Citizens would recommend to foster good IT solutions to protect their data, beyond having a strong legal framework in place.

Citizens want these conditions and concerns to be incorporated into the design of a regulatory framework that governs secondary use of data. While public support cannot be guaranteed, designing a framework based on their ideas would be an important step to building public support. The development of such a citizen powered framework does not only include appropriate governance structures but also respect for central ethical values as they are interpreted by citizens.

11. Citizens would recommend that stakeholders respect principles that align with citizens' ethical values
12. Citizens would recommend having a dynamic framework which facilitates the secondary use of health data for purposes and benefits that they support, while minimising the potential risks they identify.

What does this mean for the EHDS? It seems the EHDS and its implementation at national level could address more specifically citizens' needs to be informed about the secondary use of health data. It could also push for determining and including the value of the common good in this overarching framework, as well as fostering a plurality of views in decision-making processes and governance. Overall, the proposal and its implementation could push for citizens to be treated as equal partners. This could ensure that key concepts like privacy, consent, control, commercialisation, the common good, purpose, etc. respect citizens' conceptualisations, values and principles. Special attention could also be given to their concerns regarding their identifiability, establishing dissuasive enough sanctions mechanisms as well as ensuring that technical safeguards are in place to protect them.

Trust is reached when all elements of a data relationship are respected and there is an assumption of good intentions. The relationship exists and continues to exist in the eyes of citizens. every actor in the health data ecosystem realizes that they are entering a data relationship with data subjects and that they share a responsibility to be respectful in all elements of this relationship. It means that data users accept that they are one cog in a bigger framework and that every change they make needs to be balanced with the other elements of the data relationship. The way to work on the data relationship to ensure that it is and remains a trusted relationship, is to consider citizens as a partner in the development of a framework for a health data space.

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 Member States and associated 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.

One of the objectives of the Joint Action is to assess citizens' perspectives on the secondary use of their health data, as well as how they wish to be engaged in the future EHDS. The purpose is to ensure a bottom-up approach where citizens' voices are considered from the start and are included in the co-construction of the ethical, legal and societal framework for the secondary use of health data in Europe. To this end, the Joint Action set up the iCitizen Work Package 8 to assess citizens' perspectives, and to develop recommendations on the involvement of patients, citizens, and their representatives in the governance framework of the EHDS. To achieve this, the Work Package 8 team adopted the methodology set out below.

First, preliminary work was conducted to identify the key topics regarding the secondary use of health data from the perspective of citizens, stakeholders and ethics. The findings were then used to inform and generate input for the design of a future citizen consultation. The preliminary work included an exploratory [literature review](#) aimed at understanding the landscape and existing literature on citizens' perceptions of health data, its use for secondary purposes and their role in health data governance. The literature review was followed by 53 stakeholder interviews with individuals with expertise or lived experience in secondary use of data and citizen engagement with data sharing from the three pilot countries' health data ecosystems (Belgium, France and the United Kingdom). The interviews provided a better understanding of citizens' experiences, perceptions and needs regarding the secondary use of health data, as well as their knowledge level, expectations and desired involvement in the governance of the secondary use of health data.

This preliminary phase informed the design and development of a public consultation called Healthy Data, which was launched on the 13th of December 2021 and was open for citizens' submissions until the 8th of May 2022. The Healthy Data consultation was conducted in collaboration with the Belgian AHEAD project, which is funded by Belspo and aims to support a better integration and valorisation of the Belgian health information ecosystem. The Healthy Data consultation platform was open to all European citizens, with contributions received primarily from the three pilot countries. The aim of the public consultation was to better understand citizens' perceptions and expectations regarding the secondary use of their health data and their role in its governance. This was realised through two participative tools with different questions, described in Annex I. A communication strategy was developed to disseminate the information materials and circulate invitations to participate in the online consultation. The table below provides the available data of the number of individuals and organisations reached during this campaign.

Table 1: Available data on the number of individuals and organisations reached during the communication campaign

Health Data Newsletter	Number of subscribers	893
External partners	Number of organisations and individuals contacted	1,066
Social media optimization	Impressions/ clicks (phase 1)	1,923,836
	Impressions/ clicks (phase 2)	11,622
Results on the Healthy Data website	Number of contributions received	5,932
	Number of contributions retained	5,722
	Number of visitors (date: 09.02.2023)	39,067*

* This number is based on the number of individuals who accessed the homepage of the Healthy Data website; however, the total figure is likely to be higher. For example, this figure does not include visitors to the interactive test page.

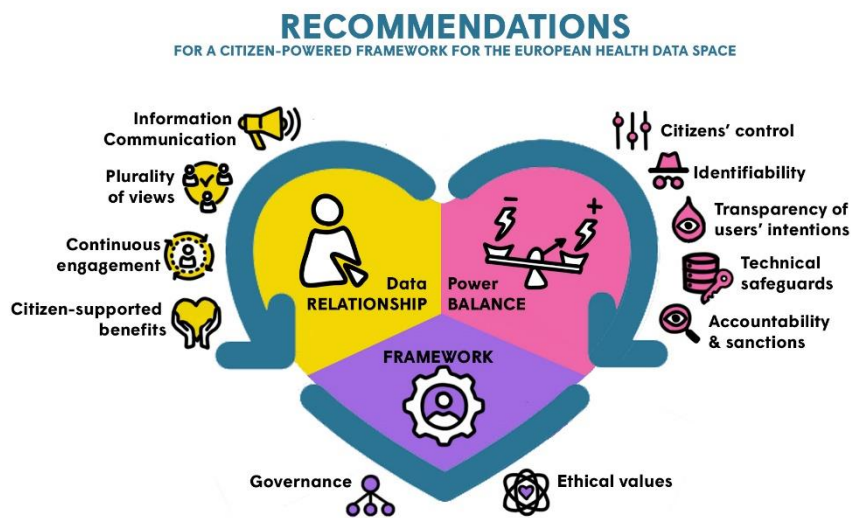
Once the consultation closed, the Work Package team proceeded to conduct a thematic analysis of the contributions gathered to produce an interim report. The aim of which was to provide an overview of key patterns and trends as well as a descriptive qualitative analysis of the citizens' opinions. This information was used as a starting point for discussion with national and European stakeholders. The interim report was published in June 2022 ([see Work Package 8 TEHDAS Consultation on TEHDAS website](#)).

Following the end of the consultation, the Work Package team organised three virtual workshops in September 2022 attended by national stakeholders from the three pilot countries. These were followed by a hybrid European workshop held in Brussels on the 11th of October 2022. During the national workshops, the Work Package team presented the results of the public consultation and gathered feedback from their national contexts. At the European workshop the Work Package team presented the results of the interim report and the outline of the citizen recommendations. The European stakeholders were then asked to discuss this outline and provide feedback. The overall aim of the four workshops was to seek stakeholders' input on how best to translate the content of the interim report (summarising citizens' views) into citizen recommendations.

This document is one of the final reports from the iCitizen Work Package (Work Package 8). It presents and describes a number of citizen recommendations, formulated to advise the European Commission and the Member States on how citizens would like to be engaged in the future EHDS. The citizen recommendations are high level, but touch on the universal concepts and discourse that ran through the Healthy Data consultation. The recommendations are each complemented by the other sources which strengthen the ethical, legal and societal analysis, namely the Healthy Data literature review, stakeholder interviews and the feedback obtained during the workshops. These recommendations are meant to represent citizens' points of view as expressed in the Healthy Data consultation and do not reflect the point of view of TEHDAS, the authors or their institutions.

3 How to read this report?

Figure 1: infographic summarising citizens' recommendations on how to engage them in the EHDS and promote a citizen powered framework.



This report is divided into four Chapters, each containing citizen recommendations. Chapter one discusses the concept of the data relationship (yellow section in Fig. 1). It explores how citizens perceive the secondary use of health data as a relationship, and how this relationship is made up of different elements that need to be considered, including how citizens want to be treated as equal partners. Chapter two stresses the need to balance the power of data (pink section in Fig. 1). Citizens perceive secondary use of health data to be powerful and beneficial but also see risks. They have different ideas on how to find a balance between these benefits and risks, which are discussed in detail in Chapter two. Finally, Chapter three consolidates the topics discussed in Chapters one and two in its discussion on the development of a citizen powered framework for the secondary use of health data (purple section in Fig. 1).

Each chapter culminates in citizen recommendations. The structure used to develop each recommendation is as follows:

- How citizens' views led us to formulate this recommendation.
- How the literature review and stakeholders' inputs echo, challenge or complete this citizen driven priority.

Finally, the implications of the citizen recommendations for the content of the current EHDS legislative proposal are explored. As the consultation was largely conducted prior to the publication of the EHDS legislative proposal, this analysis compares citizens' underlying conceptions, values and principles with those represented in the EHDS legislative proposal.

4 Recommendations

4.1 The data relationship

Every time citizens' (health) data are used, they feel that a piece of them, their identity and history, is being used, including in the case of secondary use. Therefore, it is perceived by citizens that anyone using this data is engaged in a data relationship with them. Citizens believe that they need to be respected as a partner in this relationship and that every element of the relationship -described in this Chapter- needs to be given proper attention and requires continuous work.

Citizens referred to data or types of data in many different ways. For example: my data, our data, sensitive data, valuable data, data about X or Y, data for X or Y, datasets, linked data, anonymised data, data rights, data preferences. However, in all these contexts, there was one core sentiment that was universal to all citizen contributions: "our health data belongs to us". The rationale behind these statements ('my data', 'it is our data' etc.) seems to be embedded in the recurrent conception of data by citizens as a piece of their identity, of their history, of their lives. They feel that they are *related* to it. Hence, sharing, using, and governing their health data is to enter into a data relationship with the citizen.

The citizen-ownership approach was not evident in the literature review nor during the stakeholder interviews and workshops conducted. Instead, experts suggest alternative approaches, such as state claim for public health purposes, collective ownership based on the many actors involved in the creation of the value of data or a rights-centred approach relying on a non-property relationship and the need to protect individuals' rights. In the public consultation, citizens' conception of data ownership seemed to go beyond a mere ownership or property right approach and appeared more complex. It entails some of the elements of the alternatives proposed by authors or stakeholders, such as the need to protect individual rights, the support for solidarity and secondary use directed at public health purposes, scientific progress, improving healthcare, etc.

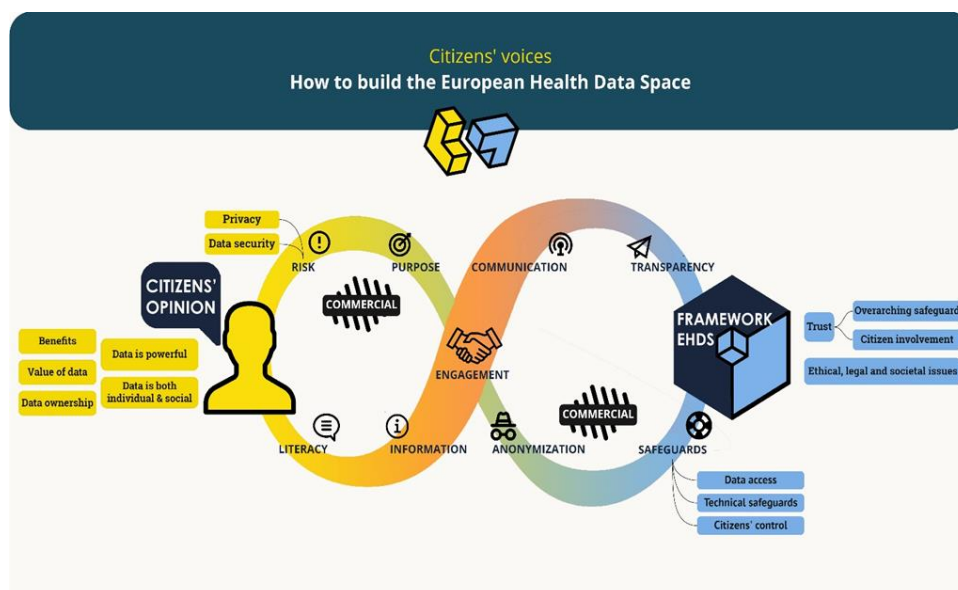
Beyond the apparent narrative of ownership, citizens recurrently shared convictions that data are about them, should benefit them and hence, that they have certain rights concerning them (e.g., to control them, be informed about their uses, etc.). They refer to themselves as the starting point and the end point of the health data cycle: the data originate from them, and the data are used to do something for/to them or to a community. Citizens perceive themselves as linked to their data across its whole cycle, even when data are linked and regardless of the format, actors, or purpose.

To illustrate citizens' conception of health data and the data relationship, consider the example of an anonymised dataset: these data are not in scope of the General Data Protection Regulation (GDPR) because they are no longer personal data. However, citizens see health data as part of themselves, a piece of their identity, whether they are anonymised or not. If these data are then used for purposes that are not supported by the citizen from whom the data were derived, they are made complicit in the achievement of this goal and feel that they are being used as a means to someone else's end. In other words: the data relationship is instrumentalised. Therefore, while the direct link with the citizen from who the data originated might be more obscure in the eye of the data user in the context of the secondary use of health

data, the position of citizens that they are the starting point and the end point in a data relationship does not change. The data originate from citizens, and they want them to do something for/to them or their community. As such, citizens are the partner with whom we need to develop a healthy data relationship.

Through the public consultation, we have been able to analyse the elements of the data relationship that were explicitly mentioned by citizens. The data relationship consists of numerous elements, all of which require proper care and attention as in any relationship. The image below presents the results from the qualitative analysis of the contributions on the Healthy Data platform. The infinity symbol interconnects the core elements of the data relationship as identified by citizens. This infinity symbol represents the complexity of the data relationship. Whenever one element changes, this can affect the balance with all other elements. The symbol also shows citizens' wish for the data relationship to flow in both directions i.e., their conceptions, values and principles should inform the framework and the framework should build a data culture that nurtures their conceptions, values and principles.

Figure 2: This image represents the data relationship and provides an overview of the main elements that citizens want to be considered in the governance framework based on the qualitative analysis of the contributions on the Healthy Data consultation platform.



Generally, when different stakeholders' conceptions and values do not match, there is a risk of 'shadowboxing': a situation where parties are never truly engaging, even if they intend to, since they hold different values or, more frequently yet more difficult to detect, they interpret values differently. This can be very dangerous because it allows for all parties to have good intentions and still fail to respect each other's rights and preferences.

For citizens, the data relationship exists, whether it is acknowledged or not. The data relationship exists between every citizen and the EHDS, not just patient and stakeholder organisations who might represent them. In this Chapter, we develop on the citizens' priorities as highlighted by the analysis of the contributions to the Healthy Data platform in order to

make recommendations on how best to take into account citizens within the data relationship and the ecosystem of secondary use of health data in Europe.

1/ Citizens would recommend being able to access information about the secondary use of health data, in an understandable way, allowing them to be more engaged

Citizens' views

Sharing data - for research - for better care

According to contributions to the public consultation, the majority of citizens support the secondary use of health data for the 'public interest', 'general interest', 'common good' etc. We refer to this as the 'common good' in this report, to distinguish it from the legal concept of the public interest. In general, this support is based on citizens' expectations that increased openness and access to health data in order to maximise its public utility will benefit patients in the future.

Strengthen educational information on the secondary use of data

Most of the contributors to the Healthy Data consultation called for more education on the secondary use of data. Some citizens admitted that they had never thought about this subject and its usefulness, did not know that data were used for research purposes or that the secondary use of data could also save lives. This led some contributors to express their lack of opinion on the sharing of health data, considering themselves insufficiently aware of its stakes and benefits as well as on its usefulness.

Nevertheless, among those who had some knowledge or visited the information materials on the website, some called for further clarification, particularly on the difference between primary and secondary use of data. Hence, we can conclude that there is a need for more education, and for raising awareness on the secondary use of health data. This educational effort was considered by citizens to be more relevant if it was based on concrete and meaningful examples and case studies on how secondary use of data works and its benefits. Some citizens expressed the need to receive understandable information on the basic concepts to enable them to understand the legal information provided. This need has been expressed by citizens as a prerequisite for trust in sharing. Moreover, it has been made clear, as mentioned in the introductory part of Chapter one, that citizens' understanding of basic concepts may be in discordance with how they are defined by other stakeholders. Examples of such concepts include data ownership, privacy, public interest, commercial interest, common good, secondary use, etc. Clarifying these concepts are essential for citizens to understand how secondary use happens, but also to ensure collective agreement on all aspects of the data relationship.

In addition, citizens have emphasised the need for information to be communicated for a public audience i.e., in a clear and easy-to-understand vocabulary and by means of accessible and comprehensible media. Visual aids such as infographics, videos or advertising posters were cited as positive examples.

For some citizens, it was also important to have confidence in the secondary use of health data and to better understand its purposes, whether the information is provided to them by

health professionals, public institutions or through their medical records. Doctors or pharmacists, for example, have been identified by citizens as trusted persons who can provide basic information on the purposes and concepts related to the secondary use of health data. Overall, informing citizens allows them to make sense of sharing health data and to engage in the secondary use of health data if they desire so. It also provides them with a sense of solidarity and contribution to research and public health improvement, and above all, it can lead to increased trust in the system.

Moving from awareness to access to further information

Some citizens expressed a desire to be able, if and when they wish, to access information on the projects that use their health data and the institutions that allow access.

Citizens mentioned the importance of receiving information and expected transparency on different elements such as: what types of data are collected and used, what types of actors use them, for what purposes, what benefits are expected and/or obtained. In order to meet this need for transparency and understanding, "portals" allowing access to the results of projects carried out with health data were suggested by citizens. This is also expressed by citizens as a prerequisite that will further strengthen trust in data sharing.

Not all citizens find their way to clear and accessible information about the rights they already have, or how to participate in, or opt out of projects where health data are used for secondary purposes. Educational information would help them to better understand and assert their rights. Moreover, this educational information was also considered important by citizens who wished to be more actively engaged. Without it, they feel less able to engage with secondary use practices.

In the context of European studies in particular, citizens also expected to know and be able to identify the competent authorities on the secondary use of health data.

Stakeholders' inputs and wider literature

Stakeholders agreed with many of the citizens' contributions, recalling the importance of providing citizens with clear and accessible information on the secondary use of health data.

Stakeholders stressed that consent procedures are sufficient and that informational materials should highlight the benefits of the secondary use of health data. Based on the analysis of individual interviews with experts within the health data ecosystem, when citizens can learn more about the sharing and secondary use of health data, and their benefits, there is a common tendency to support more data sharing for secondary purposes. Stakeholders at the European workshop also confirmed that a high level of awareness among citizens about the secondary use of health data leads to higher engagement, and that without awareness, engagement and empowerment is not possible.

It was mentioned by stakeholders that citizens' trust in the primary use of health data can be translated into trust in the secondary use, which implies that a bridge can be made in terms of information between these two types of uses. Examples were mentioned such as providing basic information on the similarities, the difference and the link between primary and secondary use of data, the difference between data for individual or societal benefits, or the differences between information and consent.

Furthermore, stakeholders attending the European workshop explained that it was more difficult for citizens to perceive the benefits of sharing their health data, compared to other data they may share, for example on social media. Therefore, the need to improve the digital literacy of citizens and to communicate shorter and more accessible information via accessible digital tools was highlighted. It is only when citizens benefit from this general digital culture and basic information that they will gradually be able to actively engage with their health data and its secondary use.

Additionally, and particularly among UK stakeholders, it was recalled that efforts should be made to involve more digitally marginalised groups and communities, referring to the digital divide that remains in terms of access to information about and engagement with health data in rural and lower socio-economic areas.

Finally, health professionals play a key role in disseminating this educational information. Stakeholders indicated that these same actors need more information and support regarding the secondary use of health data. In the same way, the experts called for awareness-raising among the actors in the field who use and manage the data.

Throughout the consultation and stakeholders' involvement, several examples can be extracted to provide accessible and pedagogical information to citizens:

- Information to make people aware of the existence of secondary use of health data and information on projects using the data could be included in the tools dedicated to primary care, such as patient health records. This information would need to be adapted to all levels of knowledge, communicated in an accessible language and in a clear and concise format. At the European level, the MyHealth@EU platform, which sets up cross-border electronic health services, could also provide information on the secondary use of health data. National or regional digital health records could also provide such information (for example Mon Espace Santé in France).
- Communication campaigns on the concept and benefits of the secondary use of health data could be put in place to provide a broader vision on the life cycle of health data, including its secondary use for research or public interest purposes. Citizens suggested formats such as advertising panels in healthcare infrastructures or videos for the general public. Content-wise, citizens and other stakeholders also recommended the build of communication material based on analogies that resonate with citizens, such as donation of blood or organs for example.
 - In France for example, the Health Data Hub, in partnership with France Assos Santé, Datalab Normandie and the University Hospital Rouen Normandy, is running a campaign to raise awareness of the secondary use of health data in Normandy, produced by design students with the aim of displaying them on hospital walls. In the UK, the Understanding Patient Data program, provides a series of easily understandable animations in different formats to explain how data saves lives.
- Health data training programmes could be run and followed voluntarily. To further elaborate the ambition of the proposed regulation to raise public awareness, Member States and associated countries should promote the development of training programs on health data or build on existing ones set up by national/ regional competent authorities. This training could be set up for the general public but also for any other interested stakeholder, to benefit from useful and accessible resources. Certain stakeholders, such as health

professionals or patient associations, constitute trusted intermediaries for citizens to improve their understanding of concepts and projects related to the secondary use of health data. Therefore, it might be appropriate to further develop the skills of health professionals and associations who wish to do so, by offering them basic training on health data. This training offer could be proposed from high school onwards and included in the education curricula as well. Such trainings could be developed collectively by different actors, including civil society and citizens, acting afterwards as ambassadors and reference trainers to disseminate these educational formats.

- In France, for example, the Health Data Hub, in partnership with France Assos Santé, the Commission nationale de l'informatique et des libertés (CNIL), the direction de la Recherche, des Études, de l'Évaluation et des Statistiques (DREES) of the French Health Ministry and the Health insurance (CNAM), has launched its first citizen training programme on the National Health Data System through pedagogical modules.

2/ Citizens would recommend having access to their data and know how they are used for secondary purposes. However, they want to choose how and when they are informed about the uses of their data.

Citizens' views

Citizens communicated their desire to be able to access information about the use of their data whenever they want. They called for accessible portals where they can search for projects that are taking place and also for projects that use their data. However, citizens' preferences for being informed about the use of their data are diverse: from those who do not want this information to reach them to those who want it to reach them without having to search for it.

Some citizens have asked for more detailed and personalised communication from organisations involved in the secondary use of health data, for example through letters and emails. Some said they were expecting high level of granularity in the types of information (types of data, use, purpose, outcomes, etc.) while others only expected general information about the use of data. For some, this will depend on how the data is used: depending on the degree of anonymisation or pseudonymisation, their needs to be informed increased or decreased. In addition, several contributions warned that an abundance of systematic and repetitive information on every project using health data would be undesirable and could be counterproductive.

Some citizens had no desire to know more about the use of their health data. This could be due, for example, to the fact that they trust the experts who use their data or simply because they are not interested. Several contributions highlighted the need to give citizens the choice to decide what information they want to receive and how they want to receive it. Contributions also showed that citizens wanted to be free to decide on the frequency of these communications.

Stakeholders' inputs and wider literature

In general, the experts stressed the importance of distinguishing between information and consent and the necessary articulation (or even conciliation) between legal information and information for the general public. They noted that citizens are also not well informed about the models for using data for research purposes (e.g., opt-out).

The experts interviewed noted that a minimum set of information should be provided as a guarantee that patients have rights and that they can exercise them. Indeed, patients should be informed about the use of their data and also about the rights they have and can claim regarding their data and their secondary use. However, individual information on the use of data cannot be uniform and information beyond this minimum should be adapted to citizens' preferences

3/ Citizens would recommend that their values should inform what is beneficial to individuals and what constitutes the common good.

Citizens' views

The importance of the purpose and the benefits intended

Citizens reminded that at all times there are individuals behind the data. These data concern their health, which they identify as a very private, sensitive, and important aspect of their lives and history. Hence, they wanted to be respected as an equal partner in this data relationship, which implies that their values are taken into account. Regarding benefits, citizens usually recognised the huge potential that the secondary use of health data can have and therefore supported it. However, beyond rejecting any potential harmful use of their data, they also shared specific conceptions of what they understood as 'beneficial'.

First, it was important for citizens that the intended benefits of using any given data were clear. Second, these purposes could be in line with and take into account citizens' values, which translates to benefiting data subjects and society's interests, as an overall principle which should guide any secondary use of health data.

The purpose and intended benefits of the secondary use of health data were very important to citizens. Several contributions highlighted the need to keep in mind and control these aspects in all the instances of the secondary use of health data: it should be defined beforehand, it should be restricted to agreed appropriate purposes, it should be a condition to access health data, and it should be controlled across the conduction of the project intended. While citizens gave many different recommendations on what information they would like to receive and how, a recurrent expressed need was to understand the benefits intended and generated by the secondary use of their data, in terms of societal benefits and, when possible, individuals. Beyond increasing awareness, this could also be a guarantee for them that health data are used for the planned and approved reason.

The purpose and the benefits allowed could be defined by citizens, for citizens

To preserve public trust, the purposes for data reuse could be guided by citizens' values. Indeed, many citizens supported the idea of using and sharing health data but had a specific idea of what purposes and benefits should be pursued when using health data. Hence, their approval was neither absolute nor unconditional.

More than being individually involved in instances of secondary use, it was more a question of being engaged as a collective voice at a higher level, where citizens are listened to and considered as a real partner when the framework is being designed. Often, citizens who expressed having no desire to be involved qualified it saying they felt their interests were already taken into account, as they trusted the current system. Others came up with concrete ideas such as developing deliberative mechanisms through public debate, consultations, more project specific engagement or continuous citizen participation. This last group also believed in the role their representatives could play, such as healthcare professionals or citizen/patient associations.

In spite of having expressed a diverse range of ideas on what purposes should be allowed, the citizens' expectations focused around societal or individual health-related purposes, such as improving research, healthcare, prevention, health policy etc. Regarding benefits, ideas ranged from individual benefit for a patient or data subject, to societal or shared benefits. Indeed, citizens felt that secondary use of health data should be guided by pursuing values such as the common good and better public health. It is important to note here that in the consultations the citizens used the terms such as "common good" and "public interest" interchangeably. We refer to this as the 'common good' in this report, to distinguish it from the legal concept of the public interest.

Therefore, citizens, patients and society should be considered the primary beneficiaries of the secondary use of data. If other benefits are expected, such as commercial benefits, these should be shared or returned to society in one way or another. More on the ethical dimension of the secondary use of health data can be found in Recommendation n°11.

Literature review

Several studies analysed within the literature review were conducted to assess citizens' perceptions of the secondary use of health data. These studies highlighted how important the aim of the use was for citizens. They also showed the importance citizens and academics placed on informing the public about the purposes, benefits and results of the secondary use of health data and called for more transparency on these topics. These studies documented widespread citizen support for projects promoting the common good, which is echoed in the results from the Healthy Data consultation.

This literature review shows that the common good and societal benefits were already considered as a baseline that should guide the secondary use of health data. Several authors viewed the common good and societal benefit as being justified because data are collected from the public, by public entities, with public money. Authors argued that this fact counterbalanced the respect for individuals' autonomy over the sharing of their health data. However, in other papers it was noted that it remained unclear what the common good was and what can or should be labelled as societal benefits. Moreover, authors added that there was too often the presumption that any research had societal value and could then promote the common good. They saw a lack of determination of what these concepts entail. However, most papers did not develop further how to determine them.

However, many of the papers assessed recommended the involvement of patient and citizens within the decision-making processes. They mentioned a diverse range of possibilities, from consultative processes to the inclusion of citizens in the evaluation of policies. Whatever the

form of engagement recommended, all suggestions made in this context advocated for an *a priori* engagement, where citizens are listened to before or at the beginning of the decision-making process, to inform future decisions that will take their values into account. The papers also included examples of citizens engaging at research or project level, where citizens could initiate or contribute to the design of the research project. Stakeholders suggested that this approach could enhance the participation of underrepresented communities and build a sense of community, and citizens from the consultation that this could improve research itself by being patient driven.

Stakeholders' inputs

Generally, stakeholders supported the secondary use of health data and believed in the potential benefits it could generate for public health and healthcare. It was commonly agreed among stakeholders that information provided to citizens should integrate the benefits of the secondary use of health data, whether societal or individual ones, to foster public trust and to increase transparency.

Their conceptions of benefits aligned with those expressed by citizens during the consultation, i.e., that it should pursue the common good and generate societal benefits, as well as individual benefits. Furthermore, several believed incentives to share data should be focused on individual benefits, either in terms of well-being or financial incentives. Regarding the pursue of the common good, stakeholders also highlighted the lack of clarity on what it encompasses.

Several stakeholders mentioned the possibility of including citizens in the definition of the criteria, the conditions and purposes considered appropriate for secondary use. Here again, citizen engagement was conceived as a collective voice rather than an individual right in the decision-making process. Some stakeholders also suggested to involve citizens in defining the ethical dimension for the secondary use of health data, or to help defining the issues that needed to be addressed the most.

4/ Citizens would recommend that decision-making processes rely on a plurality of views and actors to increase their trustworthiness, as for them the latter depends on who is involved in these instances.

Citizens' views

This consultation highlighted trustworthiness as an important element of the data relationship. To build trust in the secondary use of data, citizens typically needed to trust the people or organisations making the decisions about the use of their data. It is worth noting that according to citizens' contributions, those they need to trust can be categorised into either those involved in the decision-making process for secondary use of health data, or those involved in the governance, through providing access to data or following up practices to ensure that they are run in an appropriate manner. An additional concern from some citizens was that frameworks are defined by homogeneous groups of experts in a process lacking in transparency. As such, a solution suggested by citizens was to diversify the range of stakeholders involved in decision-making or the overall governance.

One way in which this range of stakeholders could be diversified is in terms of competences and disciplines. Fostering a plurality of visions when building the regulation as well as working with multidisciplinary committees within data governance could, according to citizens, create a new ecosystem with better protection mechanisms for secondary use of health data, through ensuring a critical and sound approach. This could have the side effect of ensuring transparency in the process of secondary use of health data and increasing public trust in the practice through the actors who oversee it.

Additionally, citizens referred to diversifying these instances in terms of directly including citizens. Citizens were divided as to whether they should be directly involved in these structures or whether to delegate these aspects to experts. There seems to be a negative correlation between the level of trust they held in decision makers and the governance, and the level of involvement they desire. However, even in the case where the public would rely on experts or their representatives rather than being directly involved, this presupposes that they are in agreement with the values that drive the secondary use of health data, such as the common good (see Recommendation n°3).

Overall, whether citizens trusted the system or not, whether they defended a form of active involvement or not, two common trends can be highlighted: Firstly, that information and transparency about the framework and the stakeholders behind it is required, in order to understand who takes part in the data relationship as a minimum. Secondly, whether citizens chose to be involved themselves or not, they desired that the values that will drive the secondary use of health data should be in line with their own values, such as the common good (see Recommendation n°3).

Stakeholders' inputs

The need for diversifying the range of stakeholders involved in decision-making processes and data governance related to the secondary use of health data, as well as the lack of clarity on who is currently involved in these latter were also highlighted by stakeholders. There needs to be transparency on all the different elements constituting this larger data relationship and ecosystem.

5/ Citizens would recommend being given the opportunity to be involved in the lifecycle of health data, as they need to be engaged on a continuous basis. Otherwise, their relationship with data custodians and users can deteriorate.

Citizens' views

Throughout the public consultation, citizens have shown that their attitudes towards the secondary use of health data are dynamic. Contextual factors that influence how citizens feel towards the secondary use of health data include who is reusing their data, why it is being reused, how it is being reused and for what purpose.

These commonly cited contextual factors often included citizens wanting to know that health data are being used in a secure manner, that research is being conducted in what citizens see as an ethical manner and whether health data are being used by commercial companies or for commercial purposes. There was also a desire to know what the benefits of health data are on an individual and societal level.

Because of the range of factors that influence citizens' views towards the secondary use of health data the collectors, controllers and secondary users of health data cannot afford to treat citizens' attitudes towards the secondary use of health data as one static position. Citizens' views are dynamic and can be subject to change on a case-by-case basis dependent on the factors at play in the secondary use. Citizens' views towards the secondary use of health data can also be influenced by events unrelated to specific instances of secondary use, for example decreasing trust in government, misuse of non-health data by big tech companies and negative media stories.

Therefore, in order to maintain the trust of citizens towards the secondary use of health data, engagement with citizens should be treated as an ongoing process. If citizens' positions are treated as static and engagement is not conducted in an ongoing manner, a scenario can arise where citizens' views towards the secondary use of health data, its collectors, controllers and secondary users can start deteriorating.

In addition to the continuous nature of engagement, collectors, controllers and secondary users of health data also need to consider the breadth and depth of engagement that they offer to citizens. The responses to the Healthy Data consultation demonstrated that when it comes to engagement, contributors' preferences were broadly assigned to one of three groups. These are:

1. Those who want to be personally engaged
2. Those who do not want to be engaged because they choose to delegate this to trusted professionals in various institutions (e.g., ethic committees and physicians and researchers).
3. Those that do not think secondary use of data should be happening, and therefore are uninterested in engagement on the topic

For those citizens that do wish to be engaged in the secondary use of health data, preferred methods ranged from one-off, project specific and continuous methods. Citizens suggested that these might include public outreach campaigns via social media, workshops, focus groups, or inclusion in institutions decision-making architecture.

For those that do not wish to be engaged directly, their rationale included a lack of interest in the secondary use of health data use and secondary use, or the view that engagement on any level would prove to be a burden. For some it also included a perception that their relationship with healthcare professionals (commonly their General Practitioner) was sufficient. They do not feel the need for personal engagement because they trust healthcare professionals to make decisions in their best interest. For example, these professionals could include those sitting on ethics, access and audit committees located across institutions with a role in health data secondary use. It should be noted that there is also a fourth group of citizens who do not engage and therefore whose views are unknown.

Those wishing to promote the secondary use of health data will be required to understand how their population views these various methods of engagement and their key motivations for being engaged. This will give secondary users an understanding of how to most effectively engage citizens in the secondary use of health data, thus promoting transparency and building trust towards reuse.

Literature review

The concept of continuous engagement appears in the literature through a number of forms. These range from the ongoing management of secondary use requests to methods for changing the way secondary use impacts specific individuals and communities.

Engagement – Education & Awareness

The literature review highlighted a view among the academic community of the importance of education, pedagogy and acculturation as methods of engagement with the public. The literature suggested that these methods were a route to improving engagement by citizens in the governance of the secondary use of health data reuse, because it would lead to data having a true meaning to them, thus allowing them to fully grasp the implications of its use.

Specifically, the literature suggested that these efforts should focus on the types and forms of secondary sharing, citizens' rights, the collaborative nature of research, the functioning of safeguards and the public benefits of secondary use and sharing.

When discussing the information that they wish to receive, citizens' preferences ranged from brief overviews of how data has been used over a set period of time, to detailed, almost real-time breakdowns of by whom, how and why data is being reused.

This Healthy Data public consultation demonstrated that in most cases the motivations for citizens were the same as those suggested by the literature. However, whilst the information desired by citizens was similar to the information the academic community advocated for, citizens differed from the academics in the sense that they did not express a desire for this information to form part of a strategic programme of education, that is designed to bring their level of understanding of data up to the required threshold. To overcome this, respect must still be paid to citizens' desired levels of engagement and the information they wish to receive. Otherwise, education efforts risk enforcing a hierarchical flow of knowledge that undermines the original intention to empower citizens, through the use of continuous personalised engagement.

Inequalities and Marginalisation

When citizens discussed the desired purposes of secondary use of data, they presented a combination of personal and societal benefits; however, when discussing engagement with health data those stating a wish to be engaged gave reasons that were more immediately personal. For example, knowing that research is being conducted in an ethical way, that the data they have contributed has had a positive impact on research, or if it will result in changes to their personal care are all commonly given reasons in the public consultation.

The literature review was broadly in line, if slightly more developed, than citizens in this area. Whilst the literature review highlighted many of the same individual benefits of engagement it also raised the concept of distributive justice, and the evaluation of the secondary use of health data for its impact upon health inequalities and inclusion of traditionally excluded communities.

Stakeholders' inputs

The findings of the literature review, the interviews with experts in the UK, France and Belgium and the national TEHDAS workshops highlighted methods of continuous engagement as a

specific route towards overcoming the risks of worsening health inequalities, or further entrenching marginalisation for vulnerable communities.

This theme of marginalisation has continued and developed throughout the expert interviews and in the Belgian, the French and UK national TEHDAS workshop.

Participants in the UK workshop advocated for a greater focus on underserved and marginalised communities by institutions and research funders. They expressed the view that this should be achieved by building long-term relationships and networks with these communities, at the local and national level. To do this effectively it will be important for institutions to understand the barriers that are associated with some individuals and communities being prevented from engaging in data reuse.

Belgian stakeholders also highlighted the impact that the digital divide and associated low rates of digital literacy have upon some groups of citizens ability to engage in secondary use of health data. Overcoming this inequality will require a long-term commitment from responsible institutions, as overcoming the contributing factors behind this marginalisation cannot be overcome with a single action or intervention and will instead require the coordinated and continuous intervention of multiple institutions.

Beyond this discussion of continuous engagement as a means for preventing and overcoming marginalisation and health inequalities, stakeholders presented a similarly broad view on the methods for continuous engagement to those presented by citizens. They saw engagement as taking many forms to overcome limitations associated with any form of engagement and other components of the overall governance framework, such as consent.

Building upon the views of citizens, but similarly reflecting the dynamic nature of citizens preferences towards secondary use, stakeholders thought that continuous engagement should have feedback loops and regular evaluation-built in. This would be an opportunity to determine what is working for citizens and updating approaches in light of developing best practice. They reflected the need for more dedicated resources to enable such a diverse set of mechanisms, and, similarly to the literature, that engagement should be used as a means for working with citizens to define the narrative underpinning secondary use.

4.2 Power balance

Citizens conceptualise data as power. Health data are especially seen as powerful because they are often sensitive, individual information based on which it would be easy to discriminate against certain groups of patients or cause them harm or alternatively to improve their situation or to inspire progress for future generations. This led many citizens to contribute their ideas about what a health data space should look like in the form of conditions: 'yes, if...', 'in this case...', 'but in that case...', 'when X, Y and Z...', ... They were looking to maximise benefits and mitigate risks.

This exercise becomes very complex when looking at all the elements of the data relationship as interrelated (see Figure 2): when one element changes, the entire power balance is affected for citizens. All the elements are interconnected, so when one part is changed, data users need to pay attention to the balance with all the other elements. For example, when the purpose of the data use or secondary use changes, the requirements change for safeguards,

citizens' control, information and communication, privacy protection, etc. Conversely, when no consent is obtained or limited pseudonymisation is possible, only certain purposes are acceptable. This interconnectedness is true for all elements of the data relationship, but it is most prominent when commercial actors are involved. Most citizens indicate that they feel very different about the data relationship in that case, revealing that all elements of the relationship are fundamentally affected.

This chapter is dedicated to the ways in which citizens find an acceptable power balance for the secondary use of health data. While the exact point where the scales are balanced was different for many citizens, the ways in which the balance was sought was more universal. It was all about proportionality: the amount of control, the safeguards, the anonymisation, the limitations on access, the amount of information required, etc. All aspects needed to be in proportion to one another and to the intended purpose with its associated harms and benefits. The key to developing a trusted health data space is to allow citizens' values and principles to create a justified balance for the data relationship.

The key consequence of integrating the power balancing act in the data relationship is that every actor within the health data ecosystem needs to realize that when they change one aspect of the data relationship, this can impact all other elements profoundly.

6/ Citizens would recommend being provided with the opportunity for meaningful and active decision-making in the secondary use of health data, as they value the ability to exercise control

Citizens' views

Citizens have unequivocally demonstrated that they recognise the power health data have in shaping their lives, in positive and negative directions, and the role that contextual factors can have in leading this direction of travel. Citizens should be able to determine the degree of control they want over secondary use and when they want to exercise this control, have access to control mechanisms, which do not prevent the proper functioning of the European health data space.

This is because when citizens are acutely aware of the ways the secondary use of data can affect them, failing to provide them the opportunity to exercise control in ways that align with their personal preferences risks disenfranchising them. Ultimately, over a sustained period this has the potential for a growth in negative views towards the secondary use of health data. This could lead to a reduction and withdrawal of support from citizens and at its most extreme instances of citizens withdrawing of the consent required for data to be used. This presents the risk of undermining the entire health data space. UK citizens responding to the Healthy Data e-consultation referred to the attempted implementation of the General Practice Data for Planning and Research (GPDPR) in the UK, as an example of this of this potentially negative outcome. Therefore, citizens need to be able to express their preference for what this control should look like, even if that preference is for no control at all.

The public consultation has demonstrated a range of highly differing views amongst citizens concerning the control they wish to be able to exercise over the secondary use of health data. Not only did their preferences for the degree of available control differ, but so did their

preferences for how to exercise this control. Understanding how to navigate this diverse and contextually variable collection of views is where resources need to be focused.

For some citizens, this control involved being asked to consent to each secondary use of health data, regardless of who, how and why this data is being used, whilst for others it can be as simple as an annual summary of how their health data have been used and what contributions this has made. For some citizens their control did not need to be exercised through them at all and could instead be achieved through the powers given to their representatives, and/or a multidisciplinary team of healthcare professionals in various institutions tasked with ensuring secondary use is conducted in a citizen centric way. By retaining control over the secondary of health data, citizens hoped to achieve a system of that sees the risks associated with secondary use, such as security breaches and the exploitative commercialisation of data, avoided. They also intended for this control to give them the ability to steer secondary use in a direction that places citizens as the main beneficiaries of the secondary use of health data, either through direct individual benefits such as better medical treatments, or through improvements in policy making, health system design or public health.

The variation that we have seen between citizens demonstrates that there is not only tension between the motivation of secondary users of health data and citizens, but also between the citizens themselves. In expressing their preference (or lack thereof) for control over the secondary use of health data, citizens have made a personal evaluation of the balance between harms and benefits. This personal variation and its degree of variation from that of another individual may lead to conflict between personal benefits and autonomy and the social potential of the secondary use of data that many citizens also expressed as a desired condition of secondary use of health data.

The primary goal of data collectors, controllers and secondary users offering control to citizens will be to make sure that the relationship between themselves and citizens is tailored to citizens' values and principles. Achieving this relationship will require offering a level of personalisation to citizens that strikes the right balance by offering them the level of personal control they have requested. However, at the same time doing so in a manner that limits the risk of overburdening those citizens that are not prepared for this level of personal control.

As a result, those involved in secondary use should treat citizens controls in the same way as the continuous engagement discussed in Recommendation n°5, as an ongoing process that is subject to continuous monitoring, review and updating. This should be delivered in a way that reflects the various ways that citizens are able to exercise control over the reuse secondary use of their health data, including consent, deliberative engagement and involvement of citizens and their representatives in the decision-making architecture of organisations.

Literature review

The literature review provided a significant body of work that promotes the role of citizens, patients and their representatives in the governance of the secondary use of health data.

Similar to the views expressed by citizens in the Healthy Data consultation, we saw a high degree of variability in the literature about what forms this could take. This could include participation in audit and access committees, lay-citizens becoming part of research teams,

or the forming of data cooperatives to make individual citizens the central decision makers in the secondary use of health data.

The literature echoed the view of citizens that the opportunity to engage in research is a method for steering research in directions that they saw as equitable, while minimising risks and building trust. The literature cited possible strategies to achieve approval from the public towards health data governance mechanisms, tools to strengthen democracy, a method for underrepresented communities to promote research they wish to see and building a sense of community. However, the literature seemed to be less explicit about the direct benefits citizens can provide to the quality of research.

When it comes to how this participation will be achieved, the literature advocated for models of contribution and co-creation, both of which echo the views expressed by some citizens wishing to be engaged in research. The literature also went further and suggested the adoption of an instigation model. This instigation model would see research projects started by citizens, patients, and/or their representatives.

The literature review emphasised the complexity associated with offering citizens a well-balanced degree of control but goes further than the results of the public consultation to offer more practical mechanisms by which citizens can be included in reuse governance across the full lifecycle of the secondary use of health data.

Stakeholders' inputs

European stakeholders all agreed that citizens need to constitute a greater voice in health data reuse decision-making across different modalities. Some stakeholders advocated for the engagement of individual citizens through deliberative approaches such as citizen juries, town halls and workshops. They also advocated for citizens' inclusion on permanent institutional structures too, which would include advisory and access committees.

For other stakeholders, the view that individual citizens should be engaged was not raised as a possibility. The majority, instead, advocated for the inclusion of patient representatives in the full lifecycle of research projects, and some advocated for their inclusion on ethics committees.

One idea emerging from the French national workshop was to introduce generalist committees at the local level that can scrutinise the use of data and ensure that it is in line with the purposes that citizens support.

Some stakeholders also highlighted the need for greater level of educational engagement, which echoes the views expressed in the literature review. Despite their differing approach, both UK and French stakeholders emphasised the need to predetermine the scope that citizens and their representatives can have. This will avoid engagement as a means of citizen control becoming a tokenistic exercise that could disenfranchise citizens.

Some stakeholders also advocated for the use of deliberative methods of engaging citizens but cautioned that handing too much control to lay citizens would risk overburdening them with responsibility for technical subjects.

Stakeholders also saw the value of including citizens and/or their representative on the governance boards of data using institutions.

Finally, some stakeholders also raised the possibility of developing individual vaults where individual citizens could approve the secondary use of health data. This is in line with the preferences of the citizens expressing a desire for the greatest level of control in the public consultation.

7/ Citizens would recommend to ensure the protection of individuals' identity, which they perceive as one of the best ways to balance the harms and benefits of the secondary use of health data.

Citizens' views

There is no way to assess the prior knowledge of citizens from the consultation on the techniques of data anonymisation. It is not clear if they fully understand the techniques of anonymisation and the extent to which they are truly non-identifiable. Besides, they often drew attention to their lack of awareness regarding safeguards in place to protect their data. Thus, it is difficult to properly grasp what they meant when referring to anonymisation. Therefore, for this analysis we follow the assumption that they referred to the level of identifiability they consider appropriate.

From the consultation, the results showed the most frequent theme while pursuing the thematic analysis was 'Anonymisation'. It was noted that often citizens will use the term "anonymised data" when referring to conditions that should be in place for the secondary use of health data. Similarly, one of the risks of secondary use of health data most mentioned in the consultation was the potential harm to data subjects' privacy when using these data. For example, the possibility of being re-identified, could result in potential harm, such as an increased risk of discrimination. It is worth mentioning that some citizens considered cases where it could be useful to be re-identified, for individual purposes or benefits, or to make data more valuable, even though they would support that data should be anonymised on a rule basis. But in this context, some specified that re-identifying data subjects should follow strict rules, and happen through a watertight system, trusted third parties or with the involvement of healthcare providers in a therapeutic relationship for example.

Within the narrative of data as an asset whose power needs to be balanced, anonymisation seemed often perceived as a significant tool to minimize potential risks. On the one hand, depending on the context and the circumstances of a particular data use, citizens would call upon anonymisation to counter-balance factors they identify as potentially more harmful. For example, as they associated the involvement of commercial actors with a higher risk of abuse and of being re-identified, some indicated that they should not be identifiable precisely when this type of actors was to use health data. On the other hand, whether data would be provided under an anonymised format or not would in itself have great implications regarding the conditions under which health data can be used. One example is that when data are shared under an anonymised format, citizens tended to allow a broader and more open use of data. Another example is the apparent negative correlation between the level of identifiability and the level of engagement desired. If data are anonymised, several citizens shared a decreased desire of being fully informed about or involved in data governance. Conversely, when data is considered not anonymised or when there is the possibility of being identified, these citizens wanted to be involved actively, such as through providing their consent.

However, it should not be considered that using data that is anonymised could justify the lack of involvement of citizens. Once again it is not clear if they fully understand the techniques to de-identify data. Clear communication about the latter is needed, either to reassure them that the reuse of data is safe, or to properly assess whether they still consider anonymisation techniques to be sufficient to carry the weight when balancing the power of data.

Literature review

The literature review assessed several studies involving citizens' questionnaires or surveys about the secondary use of health data. Similar to the findings from the current consultation, citizens in these studies shared a preference for sharing data under an anonymised or pseudonymised format; as well as their fear of being re-identified. However, they also similarly supported to be re-contacted with personalised feedback and benefits from the secondary use of their health data. Conditions in place to allow the latter are not further developed, while in the public consultation it was clear that additional safeguards and mechanisms should be in place to minimize risks of being identified beyond the necessary to perceive these personal feedback or benefits.

Moreover, the lack of knowledge and need to improve transparency and communication on anonymisation techniques was also commonly reported. The literature could complete this recommendation by adding that beyond reassuring citizens on the safe secondary use of health data, it could also foster the creation of a common culture around health data.

Stakeholders' inputs

Anonymisation was a less discussed topic in stakeholder interviews or across the workshops. However, when anonymisation was referred to, stakeholders' inputs were similar to citizens' perspectives in multiple ways. The debates were focused on the same questions, whether there should be a way of re-identifying citizens or if full anonymisation was indeed possible. This reinforces the idea that clear communication about how anonymisation is performed is needed, not just for citizens, but for all stakeholders involved in the secondary use of health data. Interestingly, some interviewees and workshop participants backed up citizens' preferences regarding the need to involve citizens in some way if full anonymisation was not performed, such as through consent. The role of anonymisation was also sometimes perceived as a factor of trust.

8/ Citizens would recommend that data users' intentions should be transparent and in line with purposes citizens support, as they think some users might share citizens' values more than others

Citizens' views

Data is perceived by citizens as powerful. They tended to believe that health data has a power that should be harnessed, so its secondary use can benefit the common good through improving public health, healthcare, health-related research etc. However, they also used the narrative of power in the sense that data could be used by many actors, for many different purposes, which could impact citizens in different ways, including negative ways. This explains why citizens' beneficence towards the secondary use of health data has to be moderated and

is, eventually, conditional, according to the citizens. Conditional beneficence implies that the balance between benefits and risks changes depending on the context.

One of the factors affecting this balance which came back recurrently in the public consultation was the nature of the user. According to them, some actors' intentions were less clear or transparent than others. A clear example was the case of commercial actors. In spite of recognising certain benefits or need to involve these stakeholders, it was a common trend among citizens to share concerns about their involvement. They associated commercial actors with a higher risk of abuse and negative impact on citizens. They also frequently put into question the compatibility between pursuing both the common good and financial interest and doubt commercial actors' good intentions.

These concerns about bad intentions and potential misalignment with what citizens considered an appropriate and desirable secondary use prompted some citizens to suggest that conditions and safeguards in place should change depending on the actor involved. Citizens supported this argument because they perceived the balance being tipped towards increased risk when certain actors are involved, such as commercial actors. Therefore, citizens suggested conditions should change when this type of user is involved. For example, they mentioned different limitations that should be in place, such as the purposes for which they can access data, their involvement in some data management respects (such as data governance, data storage or data processing), or their access to only certain types of data – in that case, only anonymised data.

In essence, citizens expected certain conditions for accessing data to be in place to guarantee that the intention behind the secondary use is in line with their values. For example, a clear and legitimate purpose should be a condition to access health data, which should be followed up to ensure it is respected. Furthermore, they called for more transparency on how these conditions are fulfilled.

Another important matter was how to guarantee that benefits could be shared or returned in some way to society and individuals. There, citizens considered that a control over benefits generated should be ensured, and that some benefits should flow from these actors' access to data, such as free services, affordable treatments or publication of results. Some also referred to financial reward for individuals or payment to public services. Finally, they required information and transparency over the involvement of these actors, their intentions and the results of their use of data.

Literature review

Several studies reported findings from questionnaires or surveys directed to citizens about the secondary use of health data. Some findings were quite similar to those from this consultation. These studies show citizens' support for control mechanisms and conditions for accessing data, to avoid potential risks from happening. The literature review similarly highlights the need to evaluate the ratio between benefits and risks.

Who is using data and for what were aspects that seemed important to citizens in these studies, and on which they also already had an opinion. They tended to show more support towards public uses or public users in the realm of health, and in the case of a lack of any public health benefit they would reject unequivocally the access to data from commercial

actors. Regarding this last point, a striking similarity is the expressed doubts about the real intentions of these actors and their pursuit of the common good.

The concerns highlighted in these studies also lead respondents to require more transparency and information about the organisations accessing data and their objectives, including the benefits produced through the secondary use of health data secondary use.

Stakeholders' inputs

Stakeholders also shared that the intention of the user was important, should be transparent and submitted to a control loop, from the beginning to the end of the process. Some would see commercial actors' involvement as problematic, or would at least understand citizens' concerns, above all because of the lack of transparency on secondary use processes. If other actors did not see an incompatibility between involving private actors and pursuing the common good, they would still promote a more transparent process and the guarantee of a return to society. As we should provide information on the benefits of the secondary use of health data to citizens, this could also be an opportunity to show what the benefits are of involving private companies. More concretely, some stakeholders agreed with citizens that depending on the interests at stake, some aspects of the framework could vary, such as the requirement for consent or the access to only anonymised data. Other suggestions in several workshops were to create a label or a certificate for ethical companies.

9/ Citizens would recommend that accountability could be enhanced through transparent and stronger mechanisms

Citizens' views

Citizens conceived of data as having the potential to generate positive outcomes when used, but also that they should be kept under control, because this power can be exerted in multiple ways, by multiple actors, and even against data subjects themselves. Hence, having the ability to manage or use data for secondary purposes means holding a certain form of power. For citizens, it seemed that a certain social contract needed to be respected while using health data, in order to avoid harm such as impeaching on privacy, exploitation or discrimination. Besides other safeguards mentioned in this report, it appeared important to citizens that being involved in the secondary use of health data implied to hold a certain responsibility. Having clear accountability mechanisms could help preventing abuses or misuses of health data, ensuring individuals' rights are respected, as well as increasing the transparency on who is responsible for their appropriate use.

Besides these preventive measures, a legally founded accountability framework should also include punitive measures that could be enforced when data have been breached or misused. Citizens also often perceived a lack of clarity on the occurrence of abuses and how they are managed as well as that the current sanctions in place are too weak.

Stakeholders' inputs

Stakeholders from the workshops joined citizens in recognising the need to have strong accountability mechanisms and to share information with the public about data breaches and how they are addressed. Moreover, one point raised during the interviews performed before setting up the consultation is that transparency should be improved not only towards citizens,

but to other types of stakeholders involved within the secondary use of health data. Some stakeholders reported reluctance to process data for secondary purposes because of the lack of clarity on the framework, particularly on data protection, and the fear to make mistakes, being pursued and sanctioned.

Across the workshops, after the publication of the EHDS legislative proposal, some stakeholders also backed up citizens' perspective that sanctions are still too weak in the current framework.

Interestingly, stakeholders from the workshops mentioned the possibility of recognising a certain responsibility of citizens themselves. They specified though that this point should be discussed provided that citizens really have access to their data and can play a certain role in their management. However, it was also reported that citizens should not be the sole gatekeepers of their data's management; or that the burden could be too heavy considering the technical complexity encompassed by the secondary use of health data.

10/ Citizens would recommend to foster good IT solution to protect their data, beyond having a strong legal framework in place.

Citizens' views

Among the different risks the secondary use of health data could entail, citizens highlighted data breaches and loss of privacy as issues that raised significant concerns and fears. Hence, citizens expressed the need to keep health data secure, no matter its use, but also shared concerns about the current status quo regarding data safety. Even though they tended to assume that there are technological limits in data security, citizens believed that technology and IT solutions are one of the most reliable safeguards that should be exploited to ensure their safety and to protect their privacy. They mentioned a range of technological availabilities such as passwords, accounts, platforms, data encryption or cyber certificates, but also tools related to privacy enhancement technologies or secure research environments.

Generally, it was common among citizens to explicitly bring attention on their lack of awareness regarding data security and safeguards in place to protect their data. Clear communication about the latter could reassure them in how the secondary use of health data is happening safely.

Stakeholders' inputs

Even if this aspect was less developed during the interviews and workshops conducted, one point raised by some interviewees ahead of the consultation's launch, was the reported lack of basic IT infrastructure in some instances, for example within healthcare institutions.

4.3 A citizen powered framework

In the data relationship, citizens considered themselves the origin (health data are derived from them) and the end point (health data should be used for their individual/collective benefit). Moreover, they required a proportionate power balance, in which harms and benefits of the secondary use of health data as well as individual control and other safeguarding mechanisms are well balanced. All these ideas and concerns should be assembled in a regulatory

framework that governs secondary use in actual practice and in a way that is truly supported by citizens. The development of such a citizen powered framework does not only include appropriate governance structures but also respect for central ethical values as they are interpreted by citizens.

11/ Citizens would recommend that stakeholders respect principles that align with citizens' ethical values

Citizens' views

Citizens repeatedly stated that the governance system for secondary use should include ethical considerations. This ethical dimension was sometimes seen as a precondition for valuable and sustainable secondary use.

On the one hand, stakeholders' adherence to ethical concerns and values was considered a way to align the secondary use of health data with the values citizens personally support. This ethical need mainly originated from the conception of health data as being very powerful, not only in a personal but also in a social way. On the other hand, the implementation of ethical standards was considered an overarching safeguard that should supplement legal rules and protection. This ethical need mainly seemed to originate from the conception of health data as being personal and sensitive.

According to citizens, major ethical principles that should be respected included the realisation and support of, on the one hand, the common good and inclusion, and, on the other hand, autonomy and control. These values reflect the dual nature of health data, as being both a social asset and an individual, very personal entity. Since both conceptualisations of health data are always concurrent, yet, depending on the specific situation, in various proportions, all values mentioned above should always be respected as much as possible and every situation requires the right balance between them. This way, the appropriate configuration of these values constructs a code of conduct for stakeholders that inspires and maintains public trust in the secondary use of health data.

The common good and inclusion

Most citizens argued that the power of health data should be harnessed. The power of data is particularly recognised in a social and collective sense and within the many purposes secondary use might serve, it should mainly aim for the realisation and support of the common good. This purpose includes specific actions such as improving healthcare or fostering scientific progress. In line with the value of the common good, many citizens argued that the benefits of secondary use should be distributed equally among everyone in society. Amplifying the idea of the common good, citizens expressed the concern of (increased) health inequalities that might be brought about by the secondary use of health data. To avoid this, people should be protected against adverse consequences of abusive health data uses, such as discrimination.

A more extensive description of the common good as a central and highly valued purpose in the secondary use of health data, can be found in Recommendation n°3.

In addition to the social perspective on the outcomes and benefits of secondary use, a similar perspective was sometimes taken towards the social contribution to the secondary use of health data. This way, the preference to harness the power of health data and to support the common good led some citizens to argue for collective involvement in secondary use. This means that to some citizens from the Healthy Data consultation, contributing to secondary use was considered self-evident or recommended and health data sharing seemed to be a civic duty. This way, everyone should be included in both the input and output of secondary use. The secondary use of health data should be inclusive and collective from start to finish and the value of solidarity should apply to both the way all citizens contribute to it as well as to the way all benefit from it. While this idea of collectively contributing to the secondary use of health data and this extension of the value of inclusion was not generally shared, it clearly illustrates the potential impact of the much more broadly supported value of the common good as a central purpose for secondary use.

Considering secondary use from the social perspective did not make citizens blind to the other benefits health data might generate. Hence, many citizens recognised that secondary use cannot merely support the common good but could also realise individual level benefits (e.g., access to research results of general interest yet also of personal, clinical relevance or access to supplementary services). More generally, this concurrence of societal and individual benefits of secondary use runs parallel to the dual nature of health data as both a social and personal entity, which is described more in detail in the introduction of Chapter 1. However, the idea of individual benefits of secondary use should be clearly dissociated from *mere* commercial or financial profit. Commercial profit for data users might be unavoidable and an inherent part of, for example, a robust pharma industry yet for citizens, commercial profit should include the guarantee that benefits are also returned to data subjects, data holders, public services or society. This way, support of the common good seems to be a *conditio sine qua non*, whereas many citizens thought that commercial profit is incompatible with societal (health) values and public interest if it is a target in itself and if it results from secondary data uses that are not in line with the common good. Hence, secondary use including commercial profit is not necessarily renounced, yet it should meet high standards and should not impede the realization of the main purpose, i.e., the support of the common good. More details on the perceived tension between the common good and the pursuit of commercial benefits can be found in Recommendation n°3 and Recommendation n°8.

Autonomy and control

Health data is not only considered powerful and versatile, but also to be very personal. The combination of these characteristics creates a common awareness regarding the potential risks of secondary use. Health data can be used for many different purposes by many different actors and not all of these purposes might align with citizens' values such as the common good and collective inclusion. Citizens fear cases of secondary use that would aim for purposes they do not support, such as profiling, intelligence and surveillance, or mere commercial profit. Those cases would not be directed at societal benefits and, instead, they could result in privacy breaches or discrimination. This way, the powerful and personal nature of health data ensures that citizens not only support values such as the common good and

inclusion, but also, as overarching value, respect for autonomy and, consequently, confidentiality and control.

Many citizens claimed that sharing health data should be an individual choice. Since health data is personal, using it without consent or against citizens' personal choices and preferences would disrespect people and their (patient) rights. Many citizens expressed the requirement of (a specific type of) consent for secondary use, a preference that is often not materialised in the current situation. The ideas of personal choice and consent obviously induce a tension with the idea of solidarity and inclusion concerning the collective contribution to secondary use. This illustrates, because of the dual nature of health data, the omnipresent combination of the values of the common good, inclusion, autonomy and control, as well as the potential tension that might arise between these values and hence, the need to find the right, context dependent balance between them. In addition to citizens' preference for personal choice regarding secondary use, they also frequently referred to the potential risks they perceive in secondary use (for instance, by failing technical safeguards), which might lead to, for example, privacy breaches or discrimination. This way, some citizens considered it also recommended or mandatory to ask for consent for secondary use because they consider their privacy and rights to be at stake and, hence, because they are taking a risk when contributing to secondary use. In some contexts, for instance when commercial actors are involved, this requirement of consent was even more emphasised.

Consent is one way of respecting citizens' values and autonomy and giving them control over the secondary use of health data and the risks they want to take. Other ways are consulting citizens about their preferences and concerns regarding secondary use or allowing people to be actively engaged in the regulatory framework for secondary use, for instance by making them part of democratic debates or ethics committees that contribute to the development of the ethical dimension of this framework. Citizens expressed a notable interest to be involved in such activities and organisations and in co-constructing and tracking ethical values and dimensions within a framework for secondary use. It emphasises the observation that citizens wanted health data to be used for purposes that are in line with their own ends and values; they wanted their values to be respected and they wanted to be able to impact and control this, both in the development and implementation of an ethical framework for secondary use. Further reference to this citizen involvement in the ethical dimension of a suitable framework can be found in Chapter 2 on the benefits of secondary use and the pursuit of the common good as a commonly supported value.

Citizens' input on grounding ethical values within a framework for the secondary use of health data shows that the rather consequentialist balance between benefits and risks is complemented by a value driven balance between supporting the common good and respecting autonomy. Citizens wanted collective benefits to be distributed among everyone, yet they also wanted to have control over whether and how they deal with the associated, potential risks.

Stakeholders' input and wider literature

The common good and inclusion

Citizens' conception regarding the dual nature of health data, as both very personal and social data, is well endorsed in literature. Whereas health data cannot be distinguished from the individuals behind them, the data only gains value when aggregated with others' data. In line with this duality, the concept of health data ownership was prominently discussed in the stakeholder workshops. There was a call to dispense with the notion of ownership, since it is legally ambiguous, creates a sense of mistrust and fear, and can lead to a perception of health data as a commodity. In secondary use, health data is not seen as more personal but is considered from a collective point of view. Therefore, stakeholders thought a conceptual shift and its implementation in societal practice should be made, in which solidarity, together with the guarantees of respect for patient rights, guide the use of health data.

The idea that secondary use should support the common good, as one of the main findings of the Healthy Data consultation, was strongly confirmed in the literature. Secondary use purposes that are of general interest seemed to be unanimously considered legitimate. In the Belgian workshop, however, there was a remarkable emphasis on the importance of personal benefits and advantages of secondary use, ranging from access to research results and/or personalised health recommendations to financial compensations. Stakeholders argued for an abolition of strict and socially desirable altruism and, instead, advocated including the full value spectrum of secondary use, including personal benefits.

The value of inclusion, referring to collectively contributing to secondary use as well as to equitably distributing its benefits, was amply discussed in the UK, Belgian, and European stakeholder workshops. Three points were raised that hamper the realisation of inclusive secondary use. Firstly, there might not be a fair representation of certain population groups in secondary use, and they may be underrepresented in the data. This makes the health information landscape misleading and can worsen health inequalities. Secondly, it is challenging to engage everyone and hear everyone's voice regarding secondary use. This can be caused by socio-demographic factors, inequalities in digital literacy, disinformation, or disinterest. Potential solutions to include underrepresented populations are building long term relationships with these subpopulations and developing (temporary) systems of trusted data custodians that represent the perspectives of these subpopulations. Recommendation n°1, Recommendation n°2 and Recommendation n°5 further elaborate on these ideas on informing and engaging citizens. Thirdly, central ethical values that should guide secondary use, such as the common good and autonomy, may be subverted by cultural relativity. This way, harmonising central values can be challenging across Europe and even within countries. As a result, there might be no clear red lines that rule out certain behaviours and actions in secondary use. Potential solutions are setting up minimal values at European level or establishing intermediary, ethical bodies that are entrusted by citizens and address and evaluate the differences between ethical values.

Autonomy and control

In all stakeholder workshops, the importance was mentioned of giving citizens (more) control over the use of their health data. Whereas citizens mainly linked control to the value of autonomy, the personal, sensitive nature of health data and the risks of secondary use, stakeholders mainly considered citizen control as required for the maintenance of public trust.

In practice, they generally suggested to provide the option of an opt-out choice regarding secondary use. However, several stakeholders and the literature also focus on the challenges of consent procedures, identifying them as unnecessary (in some contexts of public interest), unfeasible (because of the complexity and unpredictability of secondary use or because of the inapplicability to new data sources), or ineffective (because of consent forms' insufficiency to genuinely inform citizens, but also because of consent bias or the risk that too many people will choose to opt out).

Overall, it was a common idea among stakeholders that the value of autonomy is not absolute and, hence, that practices regarding control, privacy protection, or consent are neither. Specific contexts, such as public health emergencies, can grant more weight to certain values and in any situation, a feasible ratio needs to be found between benefits and risks, as well as between general interests and personal interests.

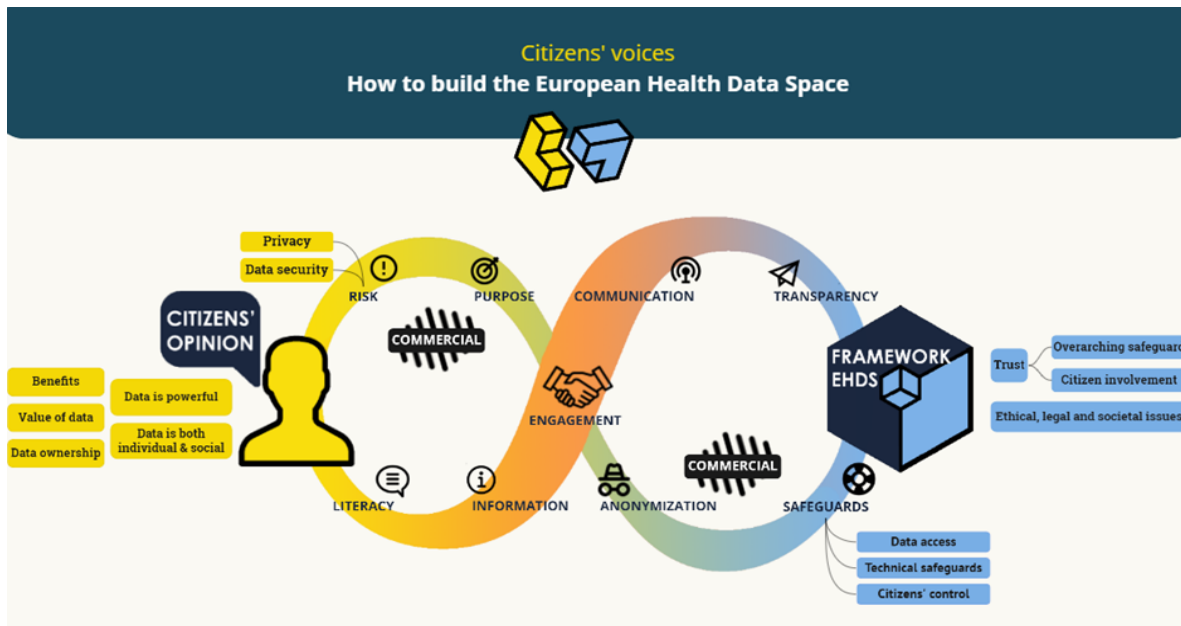
12/ Citizens would recommend having a dynamic framework which facilitates the secondary use of health data for purposes and benefits that they support, while minimising the potential risks they identify.

Citizens required their value-based ethical principles (common good and inclusion, autonomy, and control) to be translated into a strong governance framework, which ensures the practical data sharing processes are underpinned by citizens' core values.

Citizens' views

Citizens who submitted comments on an overarching framework during the Healthy Data consultation were united in their view that they should be involved in the design of the framework. In their view, the framework should apply proportionate safeguards which should be different for different purposes. They also placed high importance of the ability of the framework to facilitate the benefits of data sharing that they supported while minimising the potential risks they identified. These overarching principles underpinned the core elements of a successful framework in the eye of these citizens (see Figure 3).

Figure 3: This image represents the data relationship and provides an overview of the elements that citizens' want to be considered in the governance framework based on the qualitative analysis of the contributions on the Healthy Data consultation platform.



In order to achieve this dynamic framework, citizens wanted the framework to balance the key elements that influence their views on the secondary use of health data. These elements can be summarised as:

- Purpose: What is the data being used for, and crucially, how does this correspond to the citizens' views on what data should be used for? For example, the Healthy Data consultation showed that broadly citizens support secondary use for the common good as discussed in Recommendations n°3 and n°11.
- Benefit: What are the direct and indirect benefits of the data transfer as perceived and valued by citizens? Generally, the Healthy Data consultation showed that citizens assigned value to public benefits such as scientific progress and improvements to public health and public health organisations. Individual benefits are sometimes supported as a side effect of public benefits, whereas financial benefits are still very polarising among both citizens and stakeholders, see Recommendation n°3.
- Privacy: Can citizens be re-identified and are there any risks, as perceived by citizens, to their privacy? The contributions showed that citizens strongly welcome the EHDS principle that “privacy by design” and “bringing questions to data instead of moving the data” should be respected whenever possible. See Recommendation n°7.
- Data users: Who is using the data and can citizens trust them? Generally, the Healthy Data consultation showed that across Europe, citizens are more inclined to trust public bodies over government or commercial entities using their data. See Recommendation n°8.

Citizens want the framework to be dynamic, varying the level, number and type of safeguards required based on the balance of the elements (purpose, benefits, privacy, data users) at play in the secondary use of data. The two examples below are based on citizens' responses to specific scenarios explored during the Healthy Data Consultation and further explain what citizens meant by a “dynamic framework”.

Scenario 1: De-identified citizen health data are used by a public health organisation to support European public health research aimed at improving our understanding of treatment for diseases.

In this scenario, the data used is de-identified (see Recommendation n°7) and the research is being conducted by a public organisation in which citizens generally have a good level of trust (see Recommendation n°8). The purpose for the reuse, to improve our understanding of treatments for diseases, is generally supported by citizens (see Recommendation n°11).

The potential benefits of the research for the common good is also supported by most citizens.

Citizens' responses to this scenario showed that they were generally comfortable with the secondary use of their health data because they valued the purpose and benefits, had a level of trust in the data user and the re-identification risk was low. As a result, the citizens who responded were content with the use of anonymisation/ pseudonymisation to protect their rights and privacy during the data sharing for the purpose and benefits which align with their values.

Scenario 2: De-identified citizen health data are used by a commercial company to support health research aimed at improving our understanding of treatment for diseases.

In this scenario the data user is a commercial company. The Healthy Data consultation, corroborated by wider literature, shows that citizens generally have lower levels of trust in commercial actors using health data.

This was reflected in citizens' responses to this scenario, where they called for additional safeguards to be applied in this instance to provide higher protection against re-identification, to ensure the financial benefits for the data user are proportionate and to make sure that citizens benefit from the results of their data reuse. Without these additions, the citizens who responded were concerned about the reuse of their data in this scenario. When commercial interests were involved in the secondary use of data, citizens felt that supporting the common good was not sufficient to justify the use of their data. Citizens placed additional emphasis on the importance of autonomy and control, when considering commercial interests and the use of health data.

In summary, citizens wanted a dynamic framework, that effectively applies proportionate safeguards dependent on the 'purpose, benefit, privacy and the data users' in the secondary use of health data. In balancing these factors, the framework can facilitate a maximum-level of citizen supported data sharing, while satisfying the concerns identified by citizens.

Stakeholders' input and wider literature

Stakeholders overwhelmingly expressed their support for the creation of the EHDS and the realisation of its aims. This view that the purpose, benefit(s), re-identification risk and data user in the secondary use of health data should drive the level, type and number of safeguards, was shared by most stakeholders. Furthermore, stakeholders were particularly focused on commercial use of data, expressing both the benefits of collaborating with industry while at the same time calling for strengthened safeguards (see Recommendation n°8) and for benefit sharing i.e., ensuring that citizens benefit from the results of commercial secondary use of health data (see Recommendation n°3)

Regardless of the medium, whether that be in literature, expert interviews, national or Europe-wide workshops, stakeholders' primary comment was that citizens must be involved in the development of health data frameworks (see Recommendation n°6). However, while stakeholders universally viewed the need to involve citizens in the framework development, there were differing views on how best to engage citizens. Stakeholders saw opportunities for Europe-wide consultation and co-creation of the frameworks, policies and legislation underpinning the EHDS which would facilitate positive pan-European citizen engagement on the secondary use of health data in the EHDS.

4.4 What could this mean for the EHDS?

The results of the Healthy Data consultation and supporting literature review showed that there was strong citizen support for putting in place a legal and governance framework for the secondary use of health data at European level.

Furthermore, citizens believed that this framework, including its specifications and supporting standards, should be proportionate – balancing the benefits and risks associated with secondary use of health data. This principle is reflected in the EHDS legislative proposal which states that the initiative seeks to create an enabling framework that does not go beyond what is necessary to achieve the objectives. In citizens' view, this proportionality should be practiced through the lens of a dynamic framework which applies proportionate safeguards to balance the key factors that citizens identified: purpose, benefits, privacy, nature of the data user, etc. A dynamic framework facilitates maximum citizen support for data sharing providing a strong foundation for sustainable secondary use of health data.

This final part of the report builds on the key aspects raised by citizens in the Healthy Data consultation. However, secondary use of data is a relationship between citizens and those making decisions about the data, resulting in a complex network of interactions where citizens' views must be considered in the broader context of stakeholders' views, expert opinions, societal needs and existing practices. Citizens themselves would often refer to the need to involve experts in decision making processes about the secondary use of health data. Therefore, their views should be respected, and considered, but not necessarily always accepted as gospel.

Informing citizens and raising awareness in an educational way, about the secondary use of health data

Point to consider: Citizens indicated that the creation of the EHDS presents an opportunity to cement the need for continuous education, outreach and information strategies about the secondary use of health data in EU and national law. This role is attributed to HDABs in the current EHDS proposal, so the EU and all members states should ensure that they are adequately funded for these activities.

The European Commission's proposal for the EHDS states that HDABs should provide information to individuals about the conditions for secondary use of health data. Stakeholders welcomed these plans as measures taken to protect the rights of citizens. Beyond enabling them to exercise their rights, citizens expressed the views that information on the secondary

use of health data should fulfil an informative and awareness raising role. In their views, the information should be communicated in simple and accessible language, to build awareness and to help create a health data culture that empowers citizens to understand or get involved in the secondary use of health data.

Fostering the dissemination of information was particularly supported by citizens in the sense that they have an idea of what type of information regarding secondary use they are interested in. Citizens were interested in receiving information on the use of data itself and the impact resulting from the use of their health data. More specifically, citizens and stakeholders communicated their desire to receive information on the purpose of the secondary use of data, who was accessing the data, anonymisation and other safeguards. Citizens also called strongly for organisations to be required to state what benefits were generated from the research.

While the EHDS legislative proposal is clear on the information that must be provided on the conditions for secondary use, it is less explicit on the information which would be provided on the data applicants and users. From a citizen perspective, it was deemed that it would be very useful to have a better understanding of who is part of the relationship arising from data sharing and secondary use. To citizens, this was particularly important concerning commercial actors, and contributors made specific requests to receive information on how these actors were involved in the secondary use of health data.

The results of the Healthy Data consultation showed that opinion is divided on commercial access and use of health data. While many citizens and stakeholders support commercial entities such as pharmaceutical companies' involvement in the secondary use of health data, there was also strong evidence that many citizens and stakeholders also have concerns about their involvement. To mitigate these concerns, they often suggested that different conditions should be in place for commercial actors to access health data. They called for improved transparency on commercial entities' intentions, the process through which they access health data for secondary purposes, and how the resulting benefits from the secondary use of health data are shared with or contribute to the benefit of society. The EHDS provides an exciting opportunity to improve transparency on the use of data across Europe. Stakeholders consistently noted that strong transparency is essential to public trust in the secondary use of health data both in general and for the success of the EHDS.

During the consultation citizens raised their lack of knowledge about the measures taken to protect their privacy and it seems that further clarification would be useful for citizens on the concept of anonymisation, how it is reached and, importantly, to what extent. Stakeholders consistently warned that 'full' anonymisation might not always be possible.

Finally, while citizens expressed different preferences regarding the amount and type of information they wish to receive or not, they also expressed different needs in terms of communication and access to information. Stakeholders expressed concerns that the growing trend towards the use of digital communication tools might exacerbate health inequalities and isolate individuals with lower levels of digital literacy or access. Citizens also had a very diverse range of ideas on how to receive information, in what format and how frequently. These two points will be crucial for the EHDS to consider how best to manage and address. Together citizens and stakeholders called for a framework which would ensure that these

different needs and preferences could be taken into account. In their eyes, a dynamic framework would best support their information needs and facilitate their involvement.

Determining and including the value of the common good in the secondary use of health data

Point to consider: While the EHDS proposal refers to the “public interest”, it seems that this concept does not grasp the entire scope of what citizens have in mind when they refer to “common good”. For them, the main justification for the secondary use of health data was the potential benefit for the common good, a term we chose to distinguish it from the legal concept of public interest. In their views, the common good should guide the secondary use of health data as an overall principle. In a framework for the secondary use of health data, this principle could be defined and included in line with citizens' conception of the common good –or at least through minimum standards that allow to consider cultural differences among Member States. This, including making the role of commercial actors explicit, would promote transparency and inspire trust.

One of the overarching questions in the Healthy Data consultation was “What should health data be used for?” As a result, the Healthy Data consultation has a significant body of evidence on citizens views on the different purposes for which health data should be processed for secondary use. The purposes supported by most citizens in the consultation align well with the purposes outlined in the EHDS legislative proposal.

The EHDS legislative proposal provides that health data can be processed for secondary use when done for reasons of public interest in the area of public and occupational health. Citizens also refer to the use of health data in the pursuit of ‘public interest’, but more as a general principle. Therefore, we refer to the ‘common good’ in this report to distinguish this concept from the legal notion of ‘public interest’ used in the EHDS legislative proposal. From a citizen and stakeholder perspective, the EHDS presents an opportunity to embed the value of the common good as an overarching principle guiding data access and processing and a central value in the framework for the secondary use of health data. Citizens strongly supported the common good being an *a priori* requirement which could be assessed when reviewing health data requests. Citizens frequently mentioned that the involvement of for-profit or commercial entities in the secondary use of health data could align with pursuing the common good if, in one way or another, these commercial entities share benefits with society or public entities in the realm of health. To ensure such a distribution of realised benefits, citizens referred to examples such as financial compensation, affordable treatments or access to free services.

Fostering a plurality of views within the framework on the secondary use of health data

Point to consider: Citizens are aware of the complexity of secondary use of health data. They often identify uncertainty and a need for nuance in their own perspectives about the secondary use of health data. They expect a governance structure for the EHDS that deals with this complexity by welcoming diverse actors and views to address ethical, legal and societal challenges regarding the secondary use of health data.

The EHDS legislative proposal contains various provisions to involve stakeholders including civil society and patient representatives, for example through representation in the EHDS

Board or in HDABs at national level. From a citizen perspective, the extent to which the EHDS legislative proposal encourages cooperation and exchange of best practices among Member States and associated countries and stakeholders is welcome. Indeed, some shared their concerns about a lack of heterogeneity among decision-makers and governance bodies. Citizens see the EHDS as an opportunity to push for and implement a plurality of visions at EU and national levels in high-level decision-making and governance processes, to both ensure that multiple competences are at the table and that citizens' perspectives are well represented.

Treating citizens as equal partners

Point to consider: Citizens want to be respected as the origin point of health data and the goal of secondary use of health data – i.e., to benefit them/the common good. Therefore, they demand continuous involvement in structural ways. The current EHDS proposal recognises the importance of citizen and stakeholder engagement, but it could empower citizens even more, for example by entrenching these activities in the governance structure of HDABs, while ensuring coordination and exchange of best practices at EU level.

An important point raised by citizens was the importance of being viewed as an equal partner in the data relationship ecosystem. Within the current EHDS legislative proposal, this could partly be achieved through transparency and the participation of citizen representatives in decision-making and governance throughout the secondary use of health data process.

The EHDS legislative proposal dictates transparency on secondary use of data. As seen among citizens, stakeholders and the literature, there is clear desire and impetus for citizen engagement with the secondary use of health data, resulting in much debate about how to successfully achieve the right combination of citizen control and participation mechanisms while still enabling secondary use.

From the Healthy Data consultation, we see citizens' desire to understand decision-making processes around data sharing, including who is making the decisions. Some citizens also called for their active involvement and inclusion in these decision-making processes. The EHDS legislative proposal sets out the process by which HDABs will review data use requests and based on the consultation responses, citizens are likely to be interested in understanding this process, the involved actors and whether citizen representatives might have a role in this process.

The literature, stakeholders and citizens that have been consulted for the Healthy Data project all highlighted the need for on-going engagement and the benefits of this approach. The Healthy Data consultation showed that the mechanisms by which citizens are able to exercise their rights and control should not be treated as a one-off action, nor can there be a one size fits all approach. Data collectors, controllers and users have to invest resources in understanding the changing preferences of citizens on an ongoing basis. This understanding can then be harnessed to best support citizen involvement in decisions on the secondary use of health data.

From the Healthy Data stakeholder and expert interviews, it became clear that it is a challenge for collectors, controllers and secondary users of health data to stay on top of ever-changing

best practices that develop across and within countries. Therefore, we should consider how best to coordinate these best practices at the national and international level, so that the resource requirements of developing new approaches can be minimised. For example, the EHDS legislative proposal charges the EDHS Board with a responsibility to coordinate across countries and instructs HDABs to collaborate with one another on citizen engagement. This Europe-wide coordination via the EHDS is welcome. The Healthy Data consultation suggests that currently there is a lack of coordination on citizen engagement resulting in variations in the levels of engagement with citizens across Member States and associated countries. This results in variation across populations concerning citizens' opportunities to be involved in the secondary use of health data. The EHDS provides a positive opportunity to improve collaboration in general and to facilitate a more united European approach to citizen engagement.

The Healthy Data consultation supports the EHDS decision to delegate responsibility for choosing the method of engagement to the national level. However, consideration is needed on how best to reconcile providing citizens with their desired degree of control over their health data (as requested by citizens) with the complexity associated with ensuring data are accessible and their secondary use supports the public interest. Citizens would like to be involved in this debate so that they can contribute, if they wish, to well-developed health data spaces that realise the individual and social benefits of secondary use whilst minimising the associated risks as identified by citizens.

As a result, citizens would like to have the opportunity to continue to share their views on the development of a framework for secondary use, the ethical values and governance supporting this framework. Therefore, mechanisms could be developed to support citizens' participation in the EHDS.

Identifiability of citizens

Point to consider: Anonymisation was the most discussed safeguard by citizens. While they attribute a lot of value to this concept, they lacked clarity on what it exactly means. It should be clearly communicated what the policy around identifiability is within the EHDS.

The EHDS legislative proposal states that access will be provided to data which are relevant to the intended purpose, in an anonymised format as standard, with provision in a pseudonymised format only being possible when the data user can clearly justify this. This corresponds to citizens' views that anonymisation has a striking weight in the balance between maximising benefits and minimising harms. However, when considering anonymisation it's worth noting that the EHDS legislative proposal considers the possibility for the HDAB to inform individuals or their treating health professional about a finding from a data user which may impact their health; a suggestion which was supported by citizens when reflecting on identifiability. Such a practice obviously challenges the requirement of anonymisation by default.

Sanctions and safeguards

Point to consider: Citizens feel vulnerable in the data relationship. Especially regarding secondary use of health data, they feel that data that belong to them is at stake, but they

experience a lack of knowledge and capacity to control. The use of sanctions and safeguards can reassure citizens and demonstrate the good intentions of the EHDS.

Aside from having clear accountability mechanisms, citizens advocated that these mechanisms should be accompanied with strong and dissuasive sanction mechanisms. This aligns with the current EHDS legislative proposal which includes penalties which could be enforced by HDABs, and countries are required to lay down the rules on penalties applicable to infringements of this Regulation and shall take all measures necessary to ensure that they are implemented. Accountability and sanction mechanisms were issues of vital importance for citizens in the Healthy Data consultation. Citizens also insisted on the need to rely on technical safeguards such as secure processing environments. It can be concluded that citizens welcome the EHDS legislative proposal's commitment to ensure that health data is shared in a secure processing environment.

5 Towards a trusted European Health Data Space

The existence of the Healthy Data consultation as part of the Joint Action TEHDAS, reflects the European Commission's and other stakeholders' view that citizens play an essential role in secondary use of health data and that if the vision for a EHDS is to be realised there must be public trust. Trust ultimately lies with the different members of the public, so the perceptions and values of different publics are fundamentally important.

The data relationship exists and continues to exist in the eyes of citizens. Every actor in the health data ecosystem will enter this relationship when they engage with the EHDS, so decision makers must be aware of the needs and values of citizens, requiring ongoing citizen engagement. For example, data users should recognise that they are one cog in a bigger system, and that every change they make needs to be balanced with the other elements of the data relationship. Likewise, citizens are also an important cog in this system and their views and values must be incorporated into any resulting framework.

To ensure that the EHDS becomes and remains a trusted relationship between citizens and data users, citizens must be considered a partner in the development of the framework for the European Health Data Space. The results of the Healthy Data consultation point to some of these elements and provide recommendations as a first step towards developing a toolkit to keep working on the data relationship.

Annex 1: Participative tools to the Healthy Data Consultation

I. Open discussion platform

The first participation tool was an open discussion platform directly accessible on the website. Through the creation of an account, participants could react to four key questions related to the secondary use of health data:

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

Participants could provide an answer to the question, comment or vote on answers provided by other users and share an answer on their own social networks.

II. Interactive quiz

The interactive quiz was accessible on a different website, whose link was provided on the consultation website. Participants could answer it in an anonymous way, allowing to understand the perceptions and expectations of the citizens who will not have the time or willingness to provide their contributions on the open discussion platform. It was not meant to be a quantitative survey, but rather an entertaining tool to raise their awareness on key issues as identified in our literature review.

Participants would have to provide their opinion on 9 affirmations, distributed among three short case-studies. For each affirmation, the respondent was asked to provide his opinion on a scale from “completely disagree” to “completely agree”. Each question was accompanied by an informative material. At the end of each sets of questions, and before the next case-study, an optional open box would provide respondents with the possibility to give an answer to each of the three first general questions provided in the open discussion platform. These open responses were then integrated anonymously onto the website and included the later in the analysis of contributions.

Table 2: Case-study 1

Case-study 1	Whenever we go to a doctor, hospital or pharmacy information will be collected about us and our medical history. National health services use this data to provide the best clinical care for us, but this data can also be <i>reused</i> , i.e., used for specific purposes beyond our individual care. In your opinion, what should health data be reused for?
Affirmations	<ul style="list-style-type: none"> - Data should be reused to support public health research, such as improving our understanding of prevention and treatment of diseases. - Data should also be reused for health research such as developing medicines, including where private companies and commercial purposes are involved. - Data should only be reused to provide individual patient care and ensure the essential functioning of our health care system.
Optional open box	What should your health data be reused for?

Table 3: Case-study 2

Case-study 2	You previously caught and recovered from a virus. A university research team would now like to use your health data to better understand how the virus spreads in the population in order to limit contaminations-
Affirmations	<ul style="list-style-type: none"> - Data should be reused to support public health research, such as You agree with the idea that people other than you can decide if the research team can use your information (e.g., a board of experts, an ethical committee) - I want to be informed about how the research project uses my health data (e.g., who, for which purposes, for what benefits?) - 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?
Optional open box	Under which conditions should your health data be reused? Why?

Table 4: Case-study 3

Case-study 3	A local research program has been evaluating the impact of air pollution on health. They have been using data about air pollution and health data from all residents living in your city (but only after removing any personal information that could identify residents).
Affirmations	<ul style="list-style-type: none"> - 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. - Sharing de-identified health data should be mandatory for research programmes that support the common good. - People who do not agree to provide access to their de-identified health data should pay more for their medical care.
Optional open box	How would you like to be informed and involved in the reuse of your health data?