



Towards  
European  
Health  
Data  
Space

Milestone M8.1

# **Citizens' perception of and engagement with health data secondary use and sharing in Europe – a literature review**

25 November 2021

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
Zoé Perrin	Health Data Hub
Louise Mathieu	Sciensano

### 0.2 Keywords

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

Accepted in Project Steering Group 23 November 2021.

#### 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 © 2021 TEHDAS Consortium Partners. All rights reserved. For more information on the project, please see [www.tehdas.eu](http://www.tehdas.eu).

## Contents

Executive Summary.....	3
Context .....	5
Introduction .....	6
Methodology .....	8
1 Ethical considerations on the secondary use and sharing of health data for secondary purposes and implications for citizens.....	10
1.1 The regulatory framework surrounding health data secondary use, sharing, and the related place of the citizen .....	10
1.2 Ethical considerations on data ownership .....	11
1.3 Ethical considerations regarding health data secondary use and sharing .....	12
1.3.1 The protection of individual rights .....	13
1.3.2 Societal benefits and individual interests.....	14
1.3.3 Assessing public benefits .....	15
2 Citizen's role and involvement in the governance of health data secondary use and sharing .....	17
2.1 Citizen's involvement through transparency, education, and acculturation .....	17
2.2 Citizen's involvement through consent.....	18
2.3 Citizen's involvement through the participation in decision-making and governance processes ...	19
2.4 Citizen's involvement in research projects .....	20
3 State of citizens' perceptions on health data secondary use and sharing .....	22
3.1. Citizens' relationship with health data .....	22
3.2. The conditions for trust and sharing .....	22
4 Bibliography .....	25

## List of Figures

Infographic 1 - Geographic coverage of surveys and.....	9
---	---

## List of Tables

Table 1 - Keywords used during our research on the Google Scholars database.....	8
Table 2 – Recommendations of academics on citizen's and patient's involvement mechanisms .....	21
Table 3 – Recommendations of academics on consent systems .....	21
Table 4 - Citizens' preferences with regards to consent, per publication .....	24

## Executive Summary

The secondary use and sharing of health data can be beneficial for public health and improve healthcare in multiple ways. However, many challenges of different nature arise when health data is being processed. The role of citizens in the processing of health data seems to be increasingly acknowledged at the European political level. Nonetheless, the only form of citizen's involvement that seems to exist in the legislation is the pre-requirement for consent. This exploratory literature review aims at **deepening our knowledge on citizens' perceptions of and involvement with the secondary use of health data** in Europe. It will help us identifying the key insights which have to be considered in our two next deliverables: a citizen e-consultation, and recommendations to encourage the involvement of citizens in the future European Health Data Space, which will be delivered to the EU Commission. This executive summary highlights the key conclusions that we drew from this literature review.

### Ethical considerations regarding health data secondary use and sharing

The first identified debate among our selection of articles relates to **the different definitions of the concept of data ownership**. Some authors do contest its applicability in the context on health data, and the myth that it provides absolute control to the owner. Several authors defend other approaches, such as a state claim to data ownership or to access to data, a collective data ownership, resulting from the multi-stage and collaborative process that creates the value of data, or such as the fact that the relationship between individuals and their health data would not be based on property but on the fact that data relates to them and that their use can affect their wellbeing.

The concept of ownership is usually linked to the **protection of individual rights regarding the secondary of health data**, such as the respect for autonomy and the protection of privacy of the individuals. Some authors highlight that privacy should not be presumed as an absolute right. They raise the need to evaluate the ratio between risks and benefits of a particular research. The extent to which those risks can affect us is more debated, as is the question of how to determine societal benefits. Authors suggest several criteria to evaluate the public interest of a project, including the public nature of the projected benefit, the principle of distributive justice, how benefits are shared, the need for transparency, and the extent of citizens' involvement in the secondary use of health data.

Moreover, while the respect of those rights seems to be translated in practice by asking **the explicit and informed consent of the individual**, authors often refer to it as one of the main challenges in the context of health data secondary use. Problematic aspects of this system reportedly include its lack of adaptability to technological aspects, the impossibility to anticipate all future reuses, or the impossibility to systematically come back to the person. However, not requiring consent explicitly can have, according to these authors, a negative impact on public's trust, which is according to many sources the cornerstone of secondary use and sharing of health data.

### Citizens' role and involvement

Our review highlights the relative novelty of perceiving the **citizen-patient as a key actor** of the governance of health data secondary use and sharing.

His role in the ecosystem is also debated within the academic community: while there is a consensus **on the need for greater transparency and education** with regards to health data, it is not the case for more active forms of involvement. There is for instance **no clear consensus on which kind of consent system** should be implemented to authorize the secondary use and sharing of health data. Five main types of such systems have been identified and promoted in the literature covered by this review: the opt-out system – in which health data are presumed reusable for certain purposes unless the citizen explicitly oppose to it; the traditional systematic informed consent system, which can translate in dynamic consent systems – enabling citizens to dynamically consent or oppose to the use of their related health data on a dedicated website

or application; the broad consent system – where citizens indicate at the data collection point whether they agree with the use of their related health data in the future according to certain rules or principles; the tiered consent system – defining different levels of access depending on several variables, such as the objective of the project or the nature of the user; and the meta-consent system – which enables citizens to choose which type of consent system he or she would like to use.

Beyond consent, **other forms of citizen's involvement mechanisms are also subject to debate** within the academic community. Sources covered in this review promoted different models, including the involvement of citizens in the decision-making processes related to the governance of health data, or their participation in research projects. These models varied depending on the degree of involvement and power that citizens had on the decisions and actions taken. With regards to their involvement in the governance of health data, the presented and promoted models went from the use of public consultations and surveys to inform decision-making processes at the political level, through the allowance of a seat in governance bodies or access committees of databases, to the creation of data cooperatives, in which citizens are at the centre of the decision-making process. As for their involvement in research projects the different forms identified by the literature covered contributory, co-construction and citizen-initiated models.

**It remains to be seen which methods of citizens' involvement would be the most suitable for the future European Health Data Space.** This question needs to be asked for two types of health data secondary use and sharing systems: one the one hand, systems based on data altruism, defined defined by the Data Governance Act (Article 2, paragraph 10) as data voluntarily shared by data subjects/holders including individuals or companies for general interest purposes and addressed in TEHDAS by the WP8.2; on the other hand, systems based on the reuse of health data collected by the public and private sectors.

#### State of citizens' perceptions towards health data secondary use and sharing

Finally, with regards to **citizens' perceptions** of the secondary use of health data and their governance, the literature review highlighted that they are mainly influenced by 4 factors, namely: (i) the nature and objectives pursued by actors being granted access to health data (ii) the type of governance that regulates access to health data (iii) the measures taken to ensure the confidentiality and security of the data (iv) and the level of knowledge that citizens have of the topic. According to the results of the surveys and consultations covered in this review, this knowledge is particularly low and would reportedly impede on the establishment of trust between citizens and the other stakeholders of the health data ecosystem. However, when questioned on the topic **citizens express a need for greater transparency and education** on health data.

Another insight gained from this review is that citizens are relatively **little asked to express their preferences regarding their potential involvement** in the governance of health data secondary use and sharing. Moreover, health data are perceived differently by citizens compared to other types of personal data, and they express a greater need for protection and control for the former. However, this reportedly does not have a negative impact on citizens' willingness to share. Finally, some sources of the literature also highlighted that citizens do not seem to be opposed to the cross-border sharing of their data and would even in favour of it, as long as it remains within the borders of the European Union.

## Context

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

The Joint Action has set up the Work Package iCitizen (WP8) to elaborate recommendations on the involvement of patients, citizens, and their representatives in the governance of the EHDS. A public e-consultation will be organised in late 2021 - early 2022 in three pilot countries, i.e., Belgium, France, and the United Kingdom, to better understand citizens' perceptions and expectations regarding the use of their data and their role in this governance.

To prepare this e-consultation and, more generally, to feed the reflections on EHDS, the WP8 has conducted an exploratory literature review on the perceptions and involvement of citizens in the secondary use of their related health-data in Europe.

## Introduction

Health data secondary use and sharing can be beneficial for public health and improve healthcare in multiple ways. Integrating citizens' information can improve research on health outcomes, treatments, or symptoms <sup>9 18</sup>; sharing data from and for secondary purposes can increase health knowledge<sup>7</sup> and the quality of information<sup>11</sup>, as well as decreasing study duplication, waste, and patients' exposure to avoidable harm in future research <sup>7. 11. 12. 24</sup>. The use of health data can therefore be seen as crucial to improve clinical outcomes<sup>12. 21</sup>, but is effective only when properly combined and shared <sup>12</sup>.

### What is health data secondary use?

The concept of a secondary use of health data is currently not defined in Europe: the construction of a harmonised definition between Member States is a challenge for TEHDAS' Joint Action, and its absence is an identified barrier to health data sharing<sup>83</sup>. Within the framework of TEHDAS, it is currently defined as "*any use of health data for reasons other than those for which they were collected in the first place*"<sup>84</sup>, in particular medical research, support to decision-making, and the development of innovations in the health sector. According to other definitions, it can also refer to commercial purposes<sup>25</sup>. Within the framework of health data, according to the study conducted by the Nivel Institute on the adaptation of the General Data Protection Regulation into national legislations the concept refers primarily to the use of data for policy support or scientific or historical research<sup>82</sup>.

However, many challenges arise from this type of data processing. Cross-border data sharing depends on a patchwork of national regulations. Indeed, there is still a lack of pan-European solutions to access health data for secondary purposes, as there is a lack of data sharing governance for international health research projects <sup>8.34</sup>. Health data processing can also raise privacy protection issues once the data is taken out of medical settings or when aggregated values are not sufficient and the use of individual data is required<sup>18</sup>. Those issues can also affect patient's trust <sup>9. 12</sup>. Indeed, despite the evidence that the public supports the secondary use of health data, several cases have contributed to reduce it. Examples include the lack of a coherent information campaign for the care.data programme in the UK, or the misuse of personal data in the Cambridge Analytica scandal<sup>21</sup>. The lack of trust can impact care delivery as patients may engage in risky behaviours<sup>9. 21</sup> but also the ability for users to access and use patient data for secondary purposes<sup>21</sup>. Therefore, as already shown in the past, it is important to consider citizens' concerns about the secondary use of their health data<sup>3</sup>.

The consideration of citizens and their rights with regard to health data is relatively new in the European Union. Adopted on October 24, 1995, the European Directive 95/46/EC on data protection<sup>73</sup> is the first legislation at European level to address, albeit indirectly, the role of the citizen with regards to the processing of personal - including health - data. This directive had two objectives: the harmonization of individual rights' protection measures, and the facilitation the circulation of personal data within the Single Market<sup>80</sup>. It established the specific and informed consent system as one of the five pillars on which the processing of personal data would be legitimate, and allowed the citizen to object to certain types of processing for which consent would not be required. The Directive also authorised, under appropriate safeguards, to store personal data for longer periods than what would be necessary for the purposes for which the data were collected for historical, statistical or scientific use (articles 6, 7, 8 and 14).

Adopted in 2016 and operative since 2018, the General Data Protection Regulation (GDPR) is a continuity of the 95/46/EC Directive. Consent remains a possible legal basis for the processing of personal data, but it is strengthened in the sense that its definition has been clarified (manifestation of free, specific, informed, and unambiguous consent) and that the data controller must now be able to demonstrate at any time that consent has been obtained under valid conditions. Just like the 1995 Directive, provisions on consent are accompanied by a set of rules

regarding the transparency of personal data processing, but the GDPR strengthens the legal obligations regarding the accessibility and understandability of the information provided to the citizens. Its article 12 obliges controllers to provide information in a concise, transparent, understandable, and easily accessible manner, in clear and simple terms<sup>75</sup>.

The corollary of this transparency requirement is the strengthening of individuals' rights and the facilitation of the exercise of their rights. All these measures should contribute to the protection of individuals with regards to the processing of their personal data, which is a fundamental right in EU law<sup>72</sup>.

If consent seems to be designated as a privileged mechanism of citizen's control over health data secondary use and sharing, article 6 of the GDPR provides other legal grounds that can be invoked to justify the processing of data, including the compliance with a legal obligation or the performance of a public interest task, thus opening other possibilities for the sharing and re-use of personal data<sup>75</sup>.

This exploratory literature review aims at deepening our knowledge on the relationship that links citizens to the secondary use of health data. After a description of the current regulatory framework surrounding health data secondary use and sharing, and the place of the citizens within this framework, the first part of this review presents the ethical, legal, and societal issues surrounding the secondary use of health data and citizen's involvement in this field. The second part presents examples of possible forms that could take the involvement of citizens in health data secondary use and sharing. The third part provides an overview of citizens' perceptions toward health data secondary use and sharing and identifies the conditions influencing their willingness to share their related health data.

A clarification is needed here on the terminology used in this paper to refer to the relationship of citizens with health data secondary use. The words and concepts that should be used to refer to this relationship are currently being debated in the academic community. This is more especially the case for the term "engagement". It refers to a multiform concept, which has cognitive, behavioural, affective, and institutional dimensions that differ according to the topic being studied (marketing, video games, politics, etc.). It can have several meanings depending on whether one is dealing with the subjective or objective, individual or collective dimension of the concept. At the subjective and individual level, engagement can refer to the feeling of being concerned by a topic, or to act to promote a political or moral cause. At the objective level, it can refer to consultation and political representation mechanisms through which individuals and citizens can be involved in the governance of an issue, either directly (via processes such as referenda) or indirectly (via, for example, the participation of patients' representatives in decision-making bodies). Its definition is even more complexified by the lack of definition of the term provided by the papers covered in this review.

Given the multiformity of the concept of "engagement", the choice has thus been made here to use the term "involvement" to refer to the relationship of citizens to the secondary use of health data. The meaning of involvement is twofold: it refers both to citizens' individual feeling of being directly concerned by the topic, and to the emergence of the topic of health data as a political and societal issue, on which citizens could have, if they wish so, the opportunity to express themselves and play a role in its governance<sup>1</sup>.

---

<sup>1</sup> The concept of governance is here understood according to the definition provided by the deliverable M5.8 of TEHDAS' Joint Action i.e. as "the act of governing an entity, where the entity territorially, politically, or issue-dependently, is demarcated by rule-set boundaries". TEHDAS, "M5.8: Potential health data governance mechanisms for European Health Data Space", published on September 1, 2021, p. 10.

<https://tehdas.eu/news/eu-should-rethink-policies-on-health-data-access/>.

## Methodology

This exploratory literature review is based on 71 English-published scientific articles, issued from a 20 keywords-based research on the Google Scholars' database.

**Table 1** - Keywords used during our research on the Google Scholars database

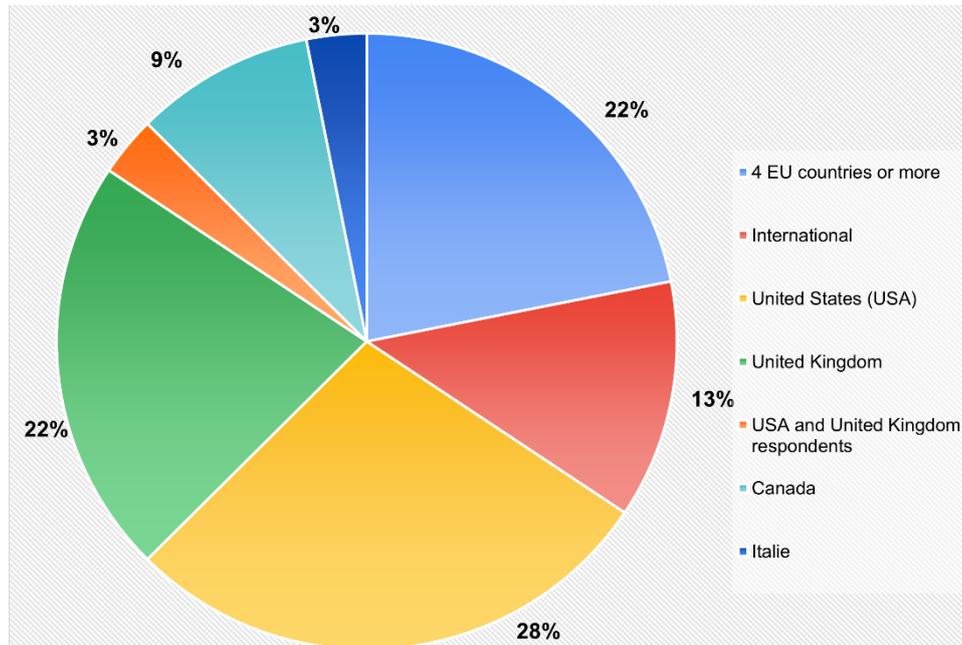
Population	Methodology	Subject	Focus
Citizen	Consultation	Secondary use	Health data
Professional	Engagement	Reuse	
Stakeholder	Survey	Data sharing	
Patients	Questionnaire		
Users	Information		
	Decision-making		
	Deliberation		
	Empowerment		
	Perception		
	ELSI		
	Sensitization		

Inclusion criteria included articles focusing on health data secondary use and/or sharing, on citizens' and stakeholders' perceptions towards this topic, on citizens' involvement in the governance of health data, and on the ethical, legal, and societal issues raised by health data secondary use and/or the involvement of citizens in its governance. Articles focusing on Europe and/or published after the promulgation of the GDPR had the priority. Exclusion criteria included publications dealing exclusively with the primary use of data (such as telemedicine, e-health, or m-health).

The scope of this literature review has several limitations and is not exhaustive. Our goal was indeed not to provide a comprehensive overview of the existing knowledge on citizens' perceptions of and involvement in health data secondary use and sharing. Some aspects of the topic of citizen's involvement in health data secondary use and sharing might be missing. The issue would thus ideally need to be explored in a more systematic and in-depth research.

Moreover, the review focuses solely on scientific articles and does not explore other sources of information, such as press articles or political publications. The selected articles are almost exclusively in English, but it would be interesting to explore the literature existing in other languages as many other articles exist on the topic. In addition, the perception studies included in this literature review focused mainly on citizens living in the United States, the United Kingdom and in Western Europe ([see infographic 1](#)). Furthermore, we conducted our research on one database, via the 'relevance' selection criterion of the Google Scholars engine, which reportedly selects sources according to the recurrence of the keywords used to perform the research, their position in the text, and the number of citations of each article<sup>77</sup>.

**Infographic 1** - Geographic coverage of surveys and consultations included in the review



# 1 Ethical considerations on the secondary use and sharing of health data for secondary purposes and implications for citizens

## 1.1 The regulatory framework surrounding health data secondary use, sharing, and the related place of the citizen

There are currently two main pieces of EU legislation regulating the processing of health data and the related place of the citizen: the Directive 2002/58/EC of the European Parliament and of the Council of July 12, 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications, which is currently being revised to become the ePrivacy Regulation) and the General Data Protection Regulation (GDPR). The EU Commission intends to complement these two texts in the coming months with several other pieces of legislation, including: the Data Governance Act, which focuses in particular on the re-use of data protected by the public sector, including health data; the upcoming legislation dedicated to the European Health Data Space; the Data Act, which aims to foster business-to-government data sharing for public interest purposes, and to support business-to-business data sharing. It will also evaluate the Intellectual Property Rights' framework to further enhance the access and use of data, and to ensure fairness in the allocation of the value of data among actors from the data economy<sup>81</sup>. The Commission's proposal on the legislation on Artificial Intelligence could also have an impact on the protection of individual rights with regards to health data processing.

The notion of health data is defined by article 4 (15) of the GDPR as all "*personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status*"<sup>75</sup>.

Health data are "*special categories of personal data*" **which benefit from a special protection regime** under article 9 the GDPR in order to protect the fundamental rights and interests of the individuals<sup>84</sup>. Health data processing is therefore being considered forbidden and can only be processed on an exceptional basis and under certain specific conditions. Consent of the individual, important public interest purposes in the field of public health or scientific, historical or statistical research are exceptions for which health data may be processed.

As mentioned in the introduction of this review, there is currently no harmonised definition of health data secondary use and sharing in Europe. Within the framework of TEHDAS' Joint Action, it is currently defined as "*any use of health data for reasons other than those for which they were collected in the first place*"<sup>84</sup>, in particular for medical research, support to decision-making, and the development of innovations in the health sector. The primary or secondary nature of health data use is irrelevant for the GDPR as long as one of the conditions of its article 9 is met. The GDPR does not favour nor restrict secondary use of data in the sense that it does not lay down separate rules for either situation, and intends to give to the data the same level of protection. The only notable difference concerns the information to be provided to the individual, which differs depending on whether the data have been collected directly from the individual (article 13) or whether it is a re-use of existing data (article 14).

The definition of health data secondary use as currently defined by TEHDAS partly matches what the GDPR calls the "*further processing*" of personal data, whose purposes must be "*compatible*" with those of the initial processing pursuant to article 5(b) of the GDPR, being understood that "*further processing for archival purposes in the public interest, for scientific or historical research purposes, or for statistical purposes shall not be considered, in accordance with article 89(1), to be incompatible with the initial purposes*".

Apart from the cases where the further processing of personal data is subject to the consent of the person or is provided for by Union law or the law of a Member State, the GDPR imposes in

its article 6(4) several criteria that need to be checked in order to determine the compatibility of the further processing, which are:

- (a) whether there is a link between the purposes for which the personal data were collected and the purposes of the further processing envisaged
- (b) the context in which the personal data were collected, in particular as regards the relationship between the data subjects and the controller
- (c) the nature of the personal data, in particular whether special categories of personal data are processed pursuant to article 9 or whether personal data relating to criminal convictions and offences are processed pursuant to article 10
- (d) the possible consequences of the proposed further processing for the data subjects;
- (e) the existence of appropriate safeguards which may include encryption or pseudonymisation.

The conditions under which health data can be processed and the applicable safeguards highlight the role played by and the protection afforded to individuals in the processing of health data, **which are based on the right of protection of natural persons** in relation to the processing of personal data. This is considered a fundamental right in European legislation<sup>72</sup>.

It is this right, and not the property right, that defines the role of citizens in relation to health data secondary use and sharing. Indeed, the property right is a legal principle that is considered as non-applicable to personal data according to the GDPR<sup>65</sup>. This does not imply that individuals have no influence on the use of health data. The right of protection of natural persons in relation to the processing of personal data induces a set of rights and principles which include, as detailed in Chapter 3 of the GDPR : the right to a transparent information – which obliges data controllers to provide information in a concise, transparent, understandable and easily accessible manner, in clear and simple terms; the right to access one's health data; the right to rectification and erasure; the right to data portability; and the right to object to the processing of data for the performance of a task carried out in the public interest, in the exercise of official authority, or for legitimate interests. However, the right to object to one's related personal data processing has limits: if the controller "*demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims*" (Chapter 3), the data controller may refuse to apply it.

This limitation of the right to object to the processing of one's related health data **illustrates the relative rather than absolute nature of the fundamental right of protection** of natural persons in relation to the processing of personal data. Recital 4 of the GDPR emphasises that it must be considered and weighed against its function in society, and balanced against other fundamental rights, in accordance with the principle of proportionality<sup>75</sup>.

Health data processing, the role that citizens play in it, and the proportionality principle applied in the application of the right of protection of natural persons raise a range of ethical, legal and political debates among the academic community. These debates will be detailed in the following paragraphs. It should be pointed out that not all the debates and ethical arguments here have a European context: some authors included in this review may have had a US, UK or international legal context when writing their articles, or even no legal reference at all. However, they are all of interest and relevance to the European Union and have therefore been included in this review.

## 1.2 Ethical considerations on data ownership

Debates on data ownership are developed in several of the selected sources<sup>2, 3, 20, 51, 67</sup>. Several authors challenge the association of data with **private property**: a feeling of ownership are sometimes existent among individuals even though they do lack a formal legal basis<sup>20</sup>. Several papers highlight that the concept is misinterpreted in the sense that ownership is often

understood as linked to the idea that who owns data has an indisputable right to control them<sup>3</sup>.<sup>20. 51</sup>. In these articles, private property rights are defined as more nuanced: several agents could reportedly claim property over health data, such as the State, through its power to acquire the state's power to acquire and use property without permission, or such as relatives in the case of genetic data<sup>3</sup>. Moreover, property rights per se do not protect the owner from non-consensual transactions<sup>20</sup>.

We can find three main arguments that challenge the conceptualisation of data as private property in the selected literature. One of them is the potential **state claim** on health data<sup>20. 51</sup>. In the case of health data generated by the public sector, this claim is justified as deriving from the public expenditures realized to generate, collect, store data and train healthcare professionals<sup>2</sup>. Other authors recall that governments have the power to use data in a non-consensual manner for legitimate public health purposes<sup>20. 51</sup>. This would not require them to claim a public ownership over health data in the end, since there would be other means to access data without the explicit consent of the related individual<sup>20</sup>.

Another debate that challenges the idea of private property is the discussion around the **creation of the value of the data**. According to some authors, designating personal data as being privately owned was justified in the past through defining personal data as characteristics that cannot be distinguished from the individual themselves<sup>45. 51</sup>. Conversely, rather than being inherent to the individual they relate to, some authors argue that the value of an individual's data does stem only once it is aggregated to others' data<sup>26</sup>, through a collaborative, multi-stage production process, where the patient provides the "raw material" that will only become clinical data once healthcare professionals will have added substantial value to it<sup>20</sup>. Its value stemming from a form of labour among different actors, there could be then a "**collective ownership**" between public (health) servants and citizens which could justify that these data are kept in the public domain<sup>3</sup>.

The third and last argument challenging the applicability of a private ownership over health data supports the idea that one can exert control without resorting to the concept of ownership. Indeed, some authors are more in favour of a "**non-property relationship**" between individuals and data. They defend the idea to switch from the idea that data belongs to us to the one that data is *about* us. As their use can affect us in many ways, individuals' rights deserve protection<sup>3</sup>. For those authors, the obligation of confidentiality of health data originates from this idea of protection rather than because we have decided to share a property.

Ownership can be considered an important question within our scope. It can determine the relationship individuals and society have with data related to them, and hence the **ways in and reasons for involving citizens** in the governance framework for health data secondary use and sharing. For instance, conceiving data as a private property is perceived as serving the idea that individuals have a right to control their data<sup>3. 20. 51</sup>. Conversely, arguing in favour of a state claim over data ownership or of the existence of a collective ownership could allow keeping data in the public domain or at least using data non-consensually for public health purposes<sup>3. 20. 51</sup>. Moreover, we can presume that defending a collective ownership over health data could result in a greater involvement of citizens, since they share ownership. Finally, even though other approaches do not reject their importance, the non-property relationship approach puts an emphasis on individual rights and their protection in the non-property relationship approach, such as the respect for autonomy and protection of privacy<sup>3</sup>.

### 1.3 Ethical considerations regarding health data secondary use and sharing

The most recurrent debate in the literature selected for this review relates to the confrontation **between the need to protect individual rights and the benefits for the general interest** which could be issued from the secondary use and sharing of health data.

### 1.3.1 The protection of individual rights

The respect of individual's autonomy and the protection of privacy are strong ethical traditions that must be considered in the governance of health data secondary use and sharing<sup>25</sup>. But in most articles included in this review, those ethical principles are not defined. Only one does provide a definition of **autonomy** as an individual freedom to live how one wants without excessive constraint from others<sup>57</sup>. We can challenge this definition in the specific case of health data as these might concern entire families<sup>45</sup>, which would imply that not only one's individual autonomy is engaged. Other papers also describe how autonomy is an evolving concept that is switching from a traditional individualistic vision of autonomy towards a more social one, where individuals reach autonomy in cooperation with others<sup>20</sup>.

Similarly, there is an assumption in a part of the literature that **privacy** could be affected in some way by health data processing, but very few papers develop the concept. One article defines it negatively as the right to not share information<sup>51</sup>, while another positively includes the right to approve the use of one's health data<sup>25</sup>. The justification behind this right could be found in Ballantyne's argumentation: data is about the patient; therefore, its use can affect their health, personal and social well-being, which is why privacy needs to be protected<sup>3</sup>. One demonstration of such protection is the confidentiality obligation, which preserves the relationship between the data and individual wellbeing within the trusting relationship with healthcare professionals<sup>3</sup>. Concerns about privacy breaches are also raised in the context of digitalization. Digitalisation challenges traditional privacy standards. These latter should thus be adapted to these new contexts, for instance by taking into consideration new sources of data, such as social networks, citizen science, or self-tracking data<sup>12, 67</sup>.

However, some authors note that privacy rights should not be presumed as absolute<sup>40</sup>, and hence there is a need to empirically evaluate the ratio between risks and benefits of a particular research<sup>21</sup>. Some concrete forms of risks in the sharing and secondary use of health data are identified in the literature. They include discrimination, stigma, the exploitation of a lack of data security<sup>2, 41</sup>. One paper focuses on the concept of the **exploitation of individuals**, and on the necessity to compare to which extent each stakeholder involved in the process is exposed when sharing data for research or development. It concludes that private entities do risk less (exposure to an investment risk) than patients, who lack control on the sharing of their related data (exposure to discrimination), are dependent on accessing healthcare systems and are obliged to share their data anyway<sup>48</sup>. Therefore, some authors agree on the necessity to implement and communicate on **accountability mechanisms**. The principle of accountability is also perceived as key to ensure public trust<sup>49</sup>, particularly in the context of the emergence of Big Data uses<sup>40, 47</sup>. Within a new paradigm where data secondary use and sharing would be promoted due to its expected societal benefits, the *a priori prohibition* of uses could be replaced by accountability mechanisms to punish data breaches and misuses<sup>40</sup>.

However, the extent to which those risks can affect our life course and experiences, and hence how to measure it, is more debated. Interestingly, it is common among different articles to compare risks from health data secondary use and sharing with material and physical harm, although they do not reach the same conclusions: some agree that data breaches are equivalent to a physical attack<sup>45</sup> (since data are equivalent to individuals themselves<sup>51</sup>); some consider that unlike interventional research, there is no body integrity invasion and that life course and experiences are not affected directly by sharing health information<sup>3</sup>; other consider that the inconvenience that may result from the secondary use of health data should not be superior to those encountered in daily life through physical and psychological routine tests<sup>57</sup>.

Furthermore, since the Second World War and the atrocities committed through scientific experiments, the traditionally key requirement to ensure the respect for autonomy and the protection of privacy in the context of biomedical research and personal data use is asking for an informed and explicit **consent**<sup>18, 34, 57</sup>. As developed in the second part of this review, the concept is mentioned in many - if not all - articles selected but is also often described as an

important challenge for health data secondary use. Indeed, some authors question its meaningfulness and its relevance for several reasons. Some mention the fact that data can refer to other individuals, such as in genetic medicine<sup>12, 45</sup>. Other authors point out that there is a possibility of consent bias because the willingness to share is usually stronger among people that are not *medically* representative of the population, i.e., people with a health condition or with an interest in health-related issues that is not representative of the general population<sup>20, 57</sup>. Some articles also highlight the existence of new private health data sources that do not systematically ask for consent when the data is extracted for research purposes, such as in social media<sup>21, 67</sup>. Other mention issues linked to technological aspects, such as the significant amount of data needed in Big Data<sup>40</sup> or that some technologies will produce unanticipated results for which consent has not been required for<sup>49</sup>. A final reason mentioned in the assessed literature is the complexity of data research projects and of their purposes, regulations, and governance, which would make it difficult to ask for an explicit and informed consent<sup>2, 12</sup>.

Nonetheless, authors recognize the importance of maintaining a certain consent model, and that increasing information can improve patients' willingness to participate in research<sup>12</sup>. Indeed, besides ensuring that they fully understand the context, asking consent is perceived as influencing patients' perspectives on research<sup>12</sup>. Not asking it may conversely have a negative impact on the trust relationship between participants and researchers for example<sup>7</sup>, which could be especially problematic as there is a large consensus in the literature on the necessity to ensure **public trust**. This latter is defined as the cornerstone of health data secondary use and sharing<sup>2, 12, 13, 20, 21, 41</sup>.

### 1.3.2 Societal benefits and individual interests

According to Porsdam et al., the primacy of the respect of autonomy is usually not debated in biomedical ethics as patients should decide what medical interventions they want to go through or not. But in the case of data processing, when it has implications for the health and wellbeing of others, and when the risks are kept to a minimum, this primacy reportedly becomes more problematic<sup>57</sup>.

This relates to another recurrent debate within the literature, which relates to finding a proper **balance between individual interests and societal benefits**<sup>11, 12, 25, 34, 47, 49</sup>. Two connected arguments justifying the importance to consider the societal benefits of health data processing can be found in the literature. First, the societal benefits that could result from data secondary use and sharing **are necessary to preserve the public good**. Authors conceptualise public good in different ways, either as public health and public health research<sup>49, 67</sup>, as science value<sup>39</sup> or as health knowledge<sup>2</sup>. Several papers also mention the "*right to science*", which is the right of all individuals "*to share in scientific advancement and its benefits*"<sup>40</sup> thanks to the secondary use and sharing of health data<sup>34, 40</sup>. A second justification is found in the argument turned to the negative that **not sharing data could be harmful for society**: potential harm can be caused "*if progress in research is not made*"<sup>21</sup> and thus, this negative impact on the healthcare system brings some authors to consider that it is reasonable for the public to expect that their data will be used to avoid it<sup>49</sup>. Hence, privacy arguments can be challenged by developing ethics discourse of the consequences of the non-use of health data<sup>21</sup>. One paper uses theories about public "*reasonable expectations*" and claims that it is reasonable to expect that our data will be used because use restrictions can have a negative impact on the healthcare system<sup>49</sup>, although one previous step here could be to identify what are the expectations that the population considers to be reasonable.

Recognising the value of the societal benefits resulting from the secondary use and sharing of health data leads some authors to consider that **explicit forms of consent could be waived** if this allows reaching such benefits<sup>49</sup> and if it cannot be obtained in a reasonable way<sup>34</sup>. Several authors go further by arguing that there could be a **moral duty to share** health information for

health research because of the public benefits that would be at stake. Hence, no requirement for any type of consent form should be implemented<sup>2, 34, 49</sup>.

But different questions arise such as **how to determine the societal benefits** of health data secondary use and sharing while protecting individual rights<sup>2</sup>. There is a need then to find a balance between all those interests at stake. If societal benefits challenge the primacy of individual interests, their emergence is not considered to be automatic for some authors. One paper denounces the presumption that is made that research has high social value. One reason for this is the difficulty to assess *a priori* how research will pay off in the end. For this author, a line needs to be drawn to identify when we can consider that secondary use or sharing health data has created societal benefits<sup>20</sup>.

A first solution is offered by Ballantyne and Schaeffer, which develop the idea of conducting a **public good test** within the framework of a moral duty to share health data<sup>2</sup>. Three major criteria can more generally be extracted from the literature covered in this review to evaluate the public interest of health data secondary use and sharing: 1/ the existence of public benefits 2/ the principle of distributive justice and 3/ the extent of citizen involvement. The ethical issues and stakes related to these questions will be explored in the following paragraphs.

### 1.3.3 Assessing public benefits

#### Public in nature?

According to Ballantyne & Schaeffer, the ethical justification for a moral duty to share health data resides both in the potential *public* ownership claim for health data collected within public infrastructures or entities and in our duty to preserve the public good of health knowledge. One may ask if such an obligation could be applied to the research conducted in private institutes or companies pursuing profitable goals, even in the case that the research results in having some social benefits<sup>2</sup>. In the same vein, other authors describe data sharing as a way of giving back the investment that they society has made in science through *public* funded research<sup>12</sup>. Similarly, Vayena et al. argue that certain data uses might be permissible for public health purposes to benefit all individuals, but not for other objectives such as corporate profit<sup>67</sup>. A last author suggests identifying “red flags” to lower the presumption that some data use offers public benefit and to determine if a use has a primarily *private* purpose<sup>20</sup>.

Therefore, it seems that some authors do perceive private/profit purposes as incompatible with public benefits. But we could ask if it is the case considering the public benefit that might result from these types of health data secondary uses or sharing. This is not further developed in the scientific papers assessed for this review. Considering that the third part of this review highlights that perceptions' studies have given a strong importance to asking participants their preferences with regards to the objectives of data processing and the nature of the users, these questions could require more ethical attention. Especially as new data sources are being owned by private entities, such as social media or data from wearables and smartphone applications, which places the individual as a “consumer” of health services<sup>41</sup>.

#### Distributive justice

Some authors suggest evaluating the social value of a data processing through the determination of **its potential benefits for the whole population**, such as the reduction of health inequity or the inclusion of populations traditionally excluded from research. Some also extend the question of the social value to the question of who will benefit from it<sup>2</sup>. One paper mentions the **principle of distributive justice**, which requires finding proportionality between the individual and common good maximization and harm minimization<sup>34</sup>.

However, authors disagree on **among whom the benefits should be shared**. Either fairness is reached if benefits are felt by those whose data have been used or shared for secondary purposes – and hence if benefits are to be felt by the whole population, there is an ethical justification for the public as a whole to bear the burden through the sharing and use of their

data for secondary purposes<sup>20</sup> - either more recent articles claim that individuals have already benefited from previous research such as through the gains that have increase life expectancy<sup>2,3</sup>.

With regards to the potential **financial incentives** that could be linked to health data secondary use and sharing, some authors claim that public willingness to benefit financially from clinical data has not been supported by empirical research<sup>3</sup>. This does however not cancel the possibility of allocating a financial value to data. Related to this aspect, some papers defend that financial gains should be directed to society as a whole and not the individuals, since the health data of one's individual become valuable only when they are aggregated with those of all the participating individuals<sup>26,48</sup>. Data would thus arguably have no real fair market value before it is used or aggregated with other health data<sup>20</sup>.

### **Citizen's involvement**

According to Ballantyne and Schaeffer, a public interest test should also include the evaluation of the degree of involvement of the public in research<sup>2</sup>. Some authors completed that argument stating that the consideration of public attitudes towards the use of health information on consent and confidentiality is often missing. While their views on data use can be positive, their level (or lack) of acceptability is reportedly assumed rather than assessed,<sup>29</sup>

As described before, some academics agree that health data secondary use and sharing may affect individual privacy rights and thus suggest that public involvement is needed to **give a voice to those affected by these policies**<sup>3</sup> or to ensure the legitimacy of the chosen governance<sup>11</sup>.

Public involvement is also described as a mechanism to enhance **public trust**<sup>2, 20, 34</sup>. Engaging the public is also seen as essential to perform *before* the secondary use and sharing of health data even happens. This would enable to define the compromise between individual rights and common good in terms of policy implementation<sup>21,49</sup>, or when research (or more broadly speaking health data secondary use and sharing considering the scope of this review) has a high social value<sup>20</sup>.

One paper develops the idea that "**cultural relativity**" could have an impact on the implementation of policies: we would reportedly need to evaluate whether privacy standards are implementable in different cultures or if they should be adapted. To gain this knowledge, we would need to assess their applicability through the understanding of individuals' vision<sup>67</sup>.

Public involvement can also reportedly increase **transparency**, since participants can spread the knowledge, they acquired on the topic<sup>11</sup>. Reciprocally, transparency can enhance public involvement: another consensus in the literature is the necessity for transparency to ensure public trust<sup>3, 7, 11, 12, 13, 41, 47, 67</sup>. Continuous communication could preserve the willingness of the citizen to be involved and minimize the risk of losing participants in research<sup>12</sup>.

## 2 Citizen's role and involvement in the governance of health data secondary use and sharing

### 2.1 Citizen's involvement through transparency, education, and acculturation

**Transparency** is promoted by a large majority of the academic community as the bedrock of citizens' involvement with their health data (*see table 2*). Beyond the guarantee that it represents in terms of control over what is being done with health data, it is also considered by a large majority of the covered literature as the pillar of any other more active roles that citizens could play in the ecosystem.

The need for transparency is particularly underlined by the authors for: the identity of the organisations being granted access to health data; the objectives being pursued; the potential risks incurred by the processing of health data; the notification of confidentiality and security breaches when they occur; and, above all, the benefits and results that are produced by health data secondary use and sharing. Several sources also promote a full transparency policy on: the decision-making mechanisms which are used for granting or denying access to health data<sup>5, 21, 39, 47, 49</sup>; existing data sharing and use agreements<sup>11, 21</sup>; and on the safeguards against the misuse of health data - in particular sanctions for misconduct <sup>10, 11, 18, 22, 57</sup>. A few authors also insist on the importance of communicating on the inherent impossibility of predicting all future processing, including when artificial intelligence and big data algorithms and systems are being involved in the process<sup>32, 49</sup>.

As stated by the GDPR, this transparency policy should according to the academic community also be based on the provision of simple, easily, accessible, and understandable information<sup>53, 55</sup>. Riggs et al promote for example the information and consent sheet template developed by the Clinical Genome Resource, based on a one-page format with a short explanatory video, as a model to ensure a clear and understandable communication<sup>59</sup>.

#### **Transparency should be combined with education, pedagogy, and acculturation efforts.**

The latter is considered as critical to ensure a proper appropriation and understanding by citizens of the topic. The creation of a common culture around health data would allow citizens not only to be informed, but also to feel directly concerned, because the data would then have a true meaning to them, enabling them to fully grasp its implications. Far from being limited to citizens, this acculturation should according to Mählmann et al. target all stakeholders of the health data ecosystem<sup>47</sup>. Educational efforts should be focused on the types and forms that health data secondary use and sharing can take; on the exercise of their rights<sup>36</sup>; on the collaborative nature of research projects that makes health data sharing necessary<sup>50</sup>; and on the functioning of safeguards and protection methods such as de-identification, anonymisation and data aggregation, which reportedly remain poorly understood <sup>10, 34, 68</sup>.

These efforts should be accompanied by **the promotion of the public benefits incurred by health data secondary use and sharing**. This is a measure that is being unanimously put forward within the academic community to induce the **perception of a direct benefit** at the individual level. Kostkova et al. suggest for instance to conduct public awareness campaigns based on empirical methods and on the provision of "success stories" <sup>41</sup>. Neves et al. also stress out the need to diffuse more efficiently the results of public perception surveys, which tend to highlight that a majority of citizens supports the secondary use and sharing of health data. This dissemination would help to deconstruct the perception of a citizen's reluctance to share that remains anchored among other stakeholders, including healthcare professionals<sup>52</sup>.

The implementation and conduct of this education and acculturation policy in practice remains however relatively unaddressed in the literature. Suggestions include encouraging a public debate around these topics, organising workshops to diffuse the information<sup>41, 49</sup>, setting up awareness-raising programmes in secondary school<sup>32</sup> or using social networks to call for action,

inform about research projects and make visible the strong support for health data sharing policies among the population<sup>27</sup>. An interesting example of such kind of initiative is Understanding Patient Data. Initially initiated by the Wellcome Trust and supported by UK institutions, it aims to create a "*national conversation*" data to improve the clarity and consistency of communication around the use of patient data. It also produces educational and awareness-raising videos to highlight the public health benefit of sharing<sup>21</sup>.

## 2.2 Citizen's involvement through consent

**Consent** seems to have so far been the main answer to the question of the implication of citizens in the governance of health data.

Three main consent paradigms emerge from the literature: **the opt-in system**, according to which health data are considered confidential and private unless the citizen expresses an explicit consent to their use and sharing; **the opt-out system**, according to which health data should be considered available for secondary use unless the citizen expresses their opposition to it; a mandatory data sharing system based **on the impossibility to opt-out system**, which is defended by a minority of the authors covered in this review. These paradigms can take a variety of forms. Kalkman et al. promote for example an opt-in paradigm that would include the possibility to use an opt-out approach when individual consent is impractical or very difficult to respect<sup>34</sup>. On the contrary, McKeown et al. argue for a restricted opt-out system, whereby citizens would be able to object to the processing of their data only for uses that do constitute a major risk to the individual<sup>49</sup>. A data sharing system based on the inability for the citizen to consent or object to the use of health data is a model that is only promoted for the health data being collected by public institutions, which could be thus considered as belonging to the public domain<sup>2</sup>.

**One can distinguish four main consent types from the literature covered in this review**, which are different implementation methods of the opt-in paradigm (*for a full picture of the recommendations provided by the literature, see table 3*).

**Informed consent**, which is the traditional consent system used in research projects, is based on the formulation of a systematic and specific agreement by the citizen-patient for any project asking to process one's related health data. This type of consent can be implemented in different ways: it can range from a classic leaflet<sup>11</sup> to the use of applications or web portals through which the citizen-patient is continuously informed of the requests formulated to process his/her related health data, and through which one can configure or adapt his or her access preferences accordingly<sup>70</sup>. This type of informed consent, also called '**dynamic consent**' or '**portable consent**', can also take the form of applications based on the principle of 'privacy-by-design', including blockchain-based systems<sup>17, 24, 45, 53</sup>. This is for example the case of MedRec, an application developed by the MIT Lab based on blockchain and "smart contract" technologies, that enables patients to access their electronic medical records and manage access authorisations<sup>51</sup>.

In **broad consent systems**, instead of expressing their informed and systematic consent, the citizen is asked at the data collection point whether they agree to see the health data being reused in the future according to certain rules or principles (for example, for public health research purposes).

**The tiered consent system** allows the citizen to define different levels of access depending on variables such as the type of health data, the identity of the user or the purpose being pursued<sup>39</sup>.

Finally, **meta-consent** systems allow individuals to choose which type of consent they would like to use in the future to express their preferences on the secondary use and sharing of their health data<sup>49</sup>.

Each of these models has its advantages and limits, one of which being the risks that a broad consent form could incur for the privacy and autonomy of the patient - i.e., the respect for each person's right to self-determination and to make choices that correspond to his or her desires, values, or life projects<sup>34</sup>. Recommendations made by the assessed literature on the involvement of citizens in health data governance seem to differ depending on the benefit/risk ratio calculated by researchers, i.e., the weight they place on the risks that sharing would incur for the privacy and autonomy of the individual, and the weight they place on its potential and effective benefits for individual and public health. This benefit/risk ratio of health data sharing seems thus to be one of the key ethical considerations surrounding health data governance.

### 2.3 Citizen's involvement through the participation in decision-making and governance processes

Often associated with the need for educational and acculturation efforts, **the involvement of patients and citizens in the decision-making process** that defines the governance framework regulating health data secondary use and sharing is also promoted by a part of the academic community as a trust-enhancer.

The regularity of this involvement and the weight allocated to citizens in the policy-making process varies across publications. A large part of the sources recommends the conduct of regular **public consultations and surveys** to inform the policy-making process, but a few authors propose more active forms of involvement mechanisms, such as the **involvement of citizens in the evaluation of public policies**<sup>47</sup> or the organisation of **citizens' juries** before any major decision. These citizen juries, which embody a more direct form of democracy, were used by Tully et al. in 2016 to discuss the creation of a new health data registry connecting citizens' electronic health records with data collected by hospitals. The two citizens' juries (34 people in total) had to decide during a 3 days-deliberation whether this register should be authorised, and, if yes, who should have access to it. The results were subsequently presented to policymakers<sup>66</sup>. This mechanism could be compared to the Climate Convention organised in 2019 in France, or the consultations on the Future of Europe announced by the Von der Leyen Commission.

A significant proportion of the literature also promotes the inclusion of citizens, patients and/or their representatives **in the governance of health databases** or national secondary use and sharing systems. This involvement method is presented as a true democratic tool, which would allow the population to fully relate to the issue, to be a true actor in the decision-making process, and which would ultimately strengthen citizens' trust in the institutions. Here again, this involvement mechanism can take several forms and degrees of intensity. It can range from a participation to advisory bodies<sup>29</sup> through the inclusion of representatives or lay citizens in decision-making committees, to the constitution of health data cooperatives in which citizens are the central decision-making actors<sup>5, 26, 47</sup>.

Among the research projects and initiatives mentioned by the academic community, the International Cancer Genome Consortium (ICGC, created in 2008), the Swiss MIDATA cooperative and the European leukodystrophy research project (LeukoTreat project) are good illustrations of this form of involvement. The LeukoTreat project is a model of involvement in which patient representatives are implicated in the development of rules that will govern the secondary use and sharing of health data. The research team involved the European Leukodystrophy Association in the development of the governance framework of the rules that will surround the use and sharing of the data collected during the study<sup>15</sup>. The ICGC corresponds to a mode of involvement of the citizen-patient which includes 'lay' people<sup>39</sup> in its independent oversight committee (the International Data Access Committee), which monitors and advises the Data Access Compliance Office, in charge of handling the access requests to the data collected by the project's members of the consortium<sup>76</sup>.

The Swiss non-profit health data cooperative MIDATA has also attracted the interest of researchers promoting a health data secondary use and sharing paradigm based on citizen empowerment. Founded in 2015 by the Eidgenössische Technische Hochschule Zürich and the Bern University of Applied Sciences, this initiative aims at facilitating health data research. It is based on the principle of the full sovereignty of the user over their related health data and the way they are used, including in an anonymised form. It is based on the MIDATA platform, which allows its users to download a copy of their data and manage access arrangements, and, in a separate space, offers access to services and smartphone applications developed by research projects. The structure includes an ethical committee elected by the general assembly of the members of the cooperative, which controls the ethical quality of the projects and services that are being offered by the research project [5.26](#).

## 2.4 Citizen's involvement in research projects

A final form method mentioned by the academic literature covered in this review relates to the **involvement of citizens in research projects**. This “citizen science” participation model is grounded in the direct impact and importance of public health in the daily life of citizens<sup>69</sup>. It would reportedly have multiple interests, including enhancing the public acceptability of the governance of health data [2.34](#), be an additional tool to strengthen the democracy of our health systems<sup>71</sup>, constitute a way for underrepresented communities to promote research projects that directly relate to them<sup>69</sup>, and contribute to building a sense of community, which increases citizens' involvement with the topic<sup>21</sup>.

Getting inspired by Wiggins and Wilbanks' classification of citizen science's models, **three main forms of research participation** can be distinguished: the contribution model, the co-construction model, and the instigation of a research project by citizens, patients, and/or their representatives.

Contributory models include crowdsourcing projects, during which citizen-patients upload self-measured health data to build up a large database and make it available for research projects. This model gained momentum with the emergence of social networks and health applications. It can be illustrated by the “Free the Data Project”, born out of the refusal of diagnostic companies to share their data. This initiative aims to make public the database of the company Myriad Genetics through the joint efforts of patients and health professionals. Another example is the American patient network PatientsLikeMe: this website is based on the premise that the data exchanged by users on the website can be used for research or commercial purposes<sup>71</sup>.

Crowdsourcing can also be used for research projects co-constructed with researchers, or even fully initiated by patients' associations or citizen groups themselves<sup>71</sup>. In these models, citizens are not only data contributors but also participate in the development, implementation and/or monitoring of the research project. Unfortunately, the literature does not mention any concrete example that specifically applies to the use of health data for secondary purposes. One can still mention the Genetics of Taste Lab Project. Initiated by the Denver Museum of Science and Nature, this project involved citizens in the collection, processing, sequencing, and analysis of genomic data voluntarily provided by museum's visitors<sup>69</sup>.

**Table 2 – Recommendations of academics on citizen's and patient's involvement mechanisms**

Involvement type	Sources
Transparency on health data secondary use and sharing practices	Aitken et al. 2016, Ballantyne & Schaeffer 2018, Blassime et al. 2018, Caine et al. 2014, Castell & Evans 2016, Colombo et al. 2019, Courbier, Dimond & Bros-Facer 2019, Cuggia & Combes 2019, Darquy et al. 2016, Despotou et al. 2020, Elger et al. 2010, Ford et al. 2019, Franklin et al. 2020, Ghafur et al. 2020, Howe et al. 2019, Jones et al. 2019, Kalkman et al. 2019 (a), Knoppers & Thorogood 2017, Knoppers & Joly 2018, Kostkova et al. 2016, Kunst et al. 2019, Lamas et al. 2015, Lemieux et al. 2021, Mählmann et al. 2017, McKeown et al. 2021, Middleton et al. 2020, Neves et al. 2019, O'Connor et al. 2017, Patil et al. 2016, Petersen 2016, Porsdam Mann et al. 2016, Riggs et al. 2019, Seltzer 2019, Spencer et al. 2019, Williams et al. 2015.
Education and sensitisation on health data secondary use and sharing	Aitken et al. 2016, Castell & Evans 2016, Colombo et al. 2019, Ghafur et al. 2020, Hassan et al. 2021, Jones et al. 2019, Karampela et al. 2019 (a,b), Knoppers & Thorogood 2017, Kostkova et al. 2016, Mählmann et al. 2017, McKeown et al. 2021, Middleton et al. 2020, Petersen 2016, Porsdam Mann et al. 2016, Shah et al. 2019, Vitak et al. 2018.
Inclusion in governance organs	Blassime et al. 2018, Colombo et al. 2019, Courbier, Dimond & Bros-Facer 2019, Cuggia & Combes 2019, De Freitas 2021, Grande et al. 2013, Hafen 2019, Hepgul et al. 2019, Kalkman et al. 2019 ab), Mählmann et al. 2017, McCoy et al. 2020, Shah et al. 2019, Winickoff et al. 2016.
Involvement in decision-making over the governance framework	Castell & Evans 2016, Colombo et al. 2019, Darquy et al. 2016, De Freitas 2021, Hay et al. 2017, Hepgul et al. 2019, Kalkman et al. 2019 (a), Kostkova et al. 2016, Kuntz et al. 2021, Jones et al. 2019, Lounsbury et al. 2021, Middleton et al. 2020, Shah et al. 2019, Tully et al. 2018.
Involvement in research projects	Ballantyne & Schaeffer 2018, De Freitas 2021, Ford et al. 2019, Hafen 2019, Kalkman et al. 2019 (a), Riggs et al. 2019, Wiggins & Wilbanks 2019, Winickoff et al. 2016.

**Table 3 – Recommendations of academics on consent systems**

Type of consent	Sources
<b>Opt-in systems</b>	
Dynamic consent	Blassime et al. 2018, Despotou et al. 2020, Lemieux et al. 2021, Pickard et al. 2014, Spencer et al 2019, Williams et al. 2015.
Specific informed consent	Colombo et al. 2019, Elger et al. 2010.
Broad consent	Courbier, Dimond & Bros-Facer 2019, Darquy et al. 2016, Jones et al. 2019, Knoppers & Thorogood 2017, Riggs et al. 2019.
Tiered consent	Caine et al. 2014, O'Connor et al. 2017, Petersen 2016
Opt-in (no precision)	Cuggia & Combes 2019, Franklin et al. 2020, Gordon & Catalini 2018, Hafen 2019, Kalkman et al. 2019 (a), Mählmann et al. 2017
<b>Opt-out system</b>	
Informed opt-out	Bell et al. 2014, Castell & Evans 2016, Ford et al. 2019, Hepgul et al. 2019, Howe et al. 2019, Kuntsman et al. 2019, McKeown et al. 2021 (restricted), Porsdam Mann et al. 2019
<b>Other systems</b>	
Meta-consent system	Cumyn et al. 2021
Need for a degree of individual control	Briscoe et al. 2020
No consent system	Ballantyne & Schaeffer 2018

### 3 State of citizens' perceptions on health data secondary use and sharing

#### 3.1. Citizens' relationship with health data

The relationship between citizens and their health data seems to be a relatively new topic in the public debate. Two major events relating to data security have contributed to the emergence of this sensitivity. First, in 2013, there was the "Snowden affair". The United States was accused of spying on citizens around the world by obtaining and analysing their smartphone data, as well as that of several large American companies. In 2018 came the "Facebook-Cambridge Analytica" scandal, involving the leakage of 87 million Facebook users' data, which was then exploited by the company Cambridge Analytica. These incidents contributed to a greater awareness of the issues related to personal data and the associated risks. They also contributed to the emergence of a public consciousness around the notions of data sharing and health data in particular<sup>23</sup>.

In 2018, the General Data Protection Regulation (GDPR), by producing a definition of personal data - and because of the obligation to implement it - also seems to have contributed to an awareness, for some citizens, of the issues related to health data<sup>78</sup>. The GDPR reinforces the specific status of health data, considered as sensitive personal data, a special status which implies that more protection is needed<sup>32</sup>.

Health data is also perceived as a category of personal data **that requires a higher protection** than other personal data according to the results of several perception studies included in this review<sup>36, 56</sup>. In contrast, the type of health data itself **does not seem to negatively impact citizens' willingness to share**. The results obtained by Middleton et al. shows, for example, that the perceived uniqueness of genomic data compared to other types of health data has either a positive or neutral relationship to willingness to share it<sup>50</sup>.

However, this growing public sensitivity to the topic is not sufficient to establish the existence of a health data culture. There is a **reported lack of knowledge of the topic** among citizens, which is mentioned in numerous publications <sup>17, 32, 50, 54</sup>. This is the case be it the issues at stake, the regulations, or the methods of collection, access, and use. Because of their relationship with the healthcare system, people with chronic diseases, long-term conditions (LTC) or rare diseases seem to be more aware of the issues related to health data<sup>12</sup>. They however constitute only a small portion of the citizens.

As indicated by the results of several studies covered by the present literature review<sup>19, 60</sup>, **citizens seem to ask for more transparency on the use of health data and** on the framework surrounding their secondary use for research<sup>4</sup>. This includes the type of processing being carried out <sup>1, 12</sup>, actors using the data, the purposes being pursued <sup>15, 22</sup>, the access procedures and regulation <sup>10, 11, 35</sup>, their rights<sup>29</sup> or the health benefits induced by the project<sup>12, 15, 22</sup>. The implementation of a health data culture seems to be even more essential as there is **reportedly a strong link between the level of knowledge** with regards to the use of health data **and the willingness to share**: the more people are informed, the more they trust and wish to share their data. However, they remain very attentive to the conditions of sharing and to their rights regarding health data<sup>50, 66</sup>.

#### 3.2. The conditions for trust and sharing

Citizens seem to be more willing to share under certain conditions. First, they are more inclined to share if the data are used for non-profit research<sup>53</sup> and if they are used by health professionals<sup>54, 56</sup>. However, the level of trust in state institutions, but also in the private sector, appears to be nuanced according to some studies<sup>54</sup>. Conversely, the fact that a public structure may be in charge seems to be an asset<sup>45</sup>.

Citizens are sensitive to the **aims of the project** and tend to be more supportive when projects are of general interest<sup>35, 36, 66</sup>. Research for the improvement of patient care and public health is thus almost unanimously considered to be a lawful and legitimate purpose for accessing, using and sharing data<sup>32</sup>. On the other hand, **if there is no benefit to public health, commercial access** to health data is unanimously **seen as unacceptable**<sup>15, 56</sup>. Citizens reportedly fear a mismatch between the stated purpose of accessing data and the real motivation of certain actors. It should be noted that actors from the **pharmaceutical industry seem to be a relatively accepted actor** by citizens despite its profit-making purpose<sup>25</sup> mainly because of the benefit brought by its activities<sup>10</sup>.

The nationality of the actors being granted access to the data does not seem to have an impact on the willingness to share at the European level: two studies clearly highlight **the willingness of citizens to share their data at European level**, recognising the benefits that cross-border data sharing can bring <sup>19, 32</sup>.

Furthermore, citizens have more confidence in health data secondary use and sharing practices when their **data are pseudonymised or even anonymised**<sup>4</sup>. Indeed, one of the main fears related to health data sharing is the risk of being re-identified<sup>70</sup>. At the same time, citizens seem to be according to some studies in favour of getting individual and personalised feedback. They want projects using their health data to provide them with advice on how to improve their health, or to inform them about the likelihood of developing a disease<sup>56</sup>.

Also, citizens are more supportive of sharing **when control mechanisms are in place and expect transparency** about the conditions of access. The fact that data is not passed on to any other organisation than the one that requested it, that there are sanctions and fines in case of misuse of the original purpose<sup>10, 35, 66</sup>, the storage of data in a secure environment<sup>10</sup>, the existence of clear and regulated access authorisation mechanisms<sup>11</sup> and the establishment of access or ethics committees<sup>15, 30, 35, 62, 66</sup> are conditions of trust and facilitate the willingness to share data.

With regards to the involvement of citizens in health data secondary use and sharing, perception studies included in this review sometimes produce contradictory results on citizens' preferences with regards to the consent system that should be implemented. This does not enable to have a clear view on this topic ([see table 4](#)). However, it seems that **citizens generally want to have a minimum degree of individual control** over what is being done with their health data. Consent systems based on weak individual control seem to be more accepted when **opt-in** is recognised as too complex or difficult to implement<sup>10, 14</sup>, if there is at least the possibility to object to the use of one's data (opt-out), if there are strong safeguards against misuse<sup>35</sup>, and if the opt-out system is accompanied by a policy of transparency<sup>35, 60, 66</sup>.

Last but not least, **the involvement of citizens in governance can reportedly also be a guarantee of trust** <sup>6, 29, 30, 63</sup>. However, only a very few studies have looked at citizens' preferences regarding their representation and involvement in governance to support this claim to make it a conclusive finding: only four of the 36 perception studies asked a question on this subject.

**Table 4 - Citizens' preferences with regards to consent, per publication**

opt-in	opt-out	No consensus identified	A (non-specified) degree of individual control is required	No question on consent
<p>Briscoe et al. 2020 Caine et al. 2014 Castell &amp; Evans 2016 Colombo et al. 2019 Courbier, Dimond &amp; Bros-Facer 2019 (citizens) Cumyn et al. 2021 Despotou et al. 2020 Pickard et al. 2014 Riggs et al. 2019 Spencer et al. 2019</p>	<p>Courbier, Dimond &amp; Bros-Facer 2019 (patients) Hepgul et al. 2019 Satinsky et al. 2018 Seltzer 2019 Tully et al. 2018</p>	<p>Aitken et al. 2016 Howe et al. 2018</p>	<p>EU Commission 2018 Jones et al. 2019 Lemieux et al. 2021</p>	<p>Bell et al. 2014 Darquy et al. 2016 Franklin et al. 2020 Grande et al. 2013 Ghafur et al. 2020 Hay et al. 2017 Karampela et al. 2019 (a) Karampela et al. 2019 (b) Kalkman et al. (b) Lounsbury et al. 2021 Middleton et al. 2020 Patil et al. 2016 Prey et al. 2016 Sarkar et al. 2020 Shah et al. 2019 Vitak et al. 2018</p>

## 4 Bibliography

### Scientific articles

1. Aitken M. et al. (2016) 'Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies', *BMC Medical Ethics* **17** (73). <https://doi.org/10.1186/s12910-016-0153-x>
2. Ballantyne A., Schaefer G. (2018) 'Consent and the ethical duty to participate in health data research', *Journal of Medical Ethics* **44**, pp. 392-396. <https://jme.bmj.com/content/44/6/392>
3. Ballantyne A. (2020) 'How should we think about clinical data ownership?', *Journal of Medical Ethics* **46**, pp. 289-294. <https://jme.bmj.com/content/46/5/289>
4. Bell, E. A., Ohno-Machado, L. and Grando, M. A. (2014) 'Sharing My Health Data: A Survey of Data Sharing Preferences of Healthy Individuals', *AMIA Annual Symposium Proceedings*, pp. 1699–1708. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4419941/>
5. Blassime A., Vayena E., Hafen E. (2018) 'Democratizing Health Research through Data Cooperatives', *Philosophy & Technology* **31**, pp. 473-479. <https://doi.org/10.1007/s13347-018-0320-8>.
6. Briscoe F., Ajunwa I., Gaddis A., McCormick J. (2020) 'Evolving public views on the value of one's DNA and expectations for genomic database governance: Results from a national survey' *PLoS ONE* **15**(3). <https://doi.org/10.1371/journal.pone.0229044>
7. Bull S. et al. (2015) 'Best Practices for Ethical Sharing of Individual-Level Health Research Data From Low- and Middle-Income Settings', *Journal of Empirical Research on Human Research Ethics* **10**(3), pp. 302-313. <https://doi.org/10.1177/1556264615594606>
8. Burgun A. et al. (2017) 'Health data for public health: toward s new ways of combining data sources to support research efforts in Europe', *Yearbook of Medical Informatics* **26**(1), pp. 235-240. <http://doi.org/10.15265/IY-2017-034>
9. Caine K., Kohn S., Lawrence C., Hanania R. (2014) 'Designing a patient-centered user interface for access decisions about HER data: implications from patient interviews', *Journal of General Internal Medicine* **30**(1), pp.7-16. <http://doi.org/10.1007/s11606-014-3049-9>
10. Castell S., Evans H. (2016). 'The One-Way Mirror: Public attitudes to commercial access to health data, Report prepared for the Wellcome Trust'. [online]. *Ipsos MORI, Social Research Institute* **155**. <https://wellcome.org/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf>
11. Colombo C., Roberto a., Krleza-Jeric K., Parmelli E. (2019) 'Sharing individual participant data from clinical studies: a cross-sectional online survey among Italian patient and citizen groups', *BMJ Open* **9**. <http://doi.org/10.1136/bmjopen-2018-024863>
12. Courbier S., Dimond R., Bros-Facer V. (2019) 'Share and protect our health data: an evidence-based approach to rare disease patients' perspectives on data sharing and data protection – quantitative survey and recommendations', *Orphanet Journal of Rare Diseases* **14**(175). <https://doi.org/10.1186/s13023-019-1123-4>
13. Cuggia M., Combes S. (2019) The French Health Data Hub and the German Medical Informatics Initiatives: Two National Projects to Promote Data Sharing in Healthcare, *Yearbook of Medical Informatics* **28**(1), 195-202. <http://doi.org/10.1055/s-0039-1677917>

14. Cumyn A. Et al. (2021) 'Citizens, Research ethics Committee Members and Researchers' Attitude Toward Information and Consent for the Secondary Use of health Data: implications for research within learning health systems', *Journal of Empirical Research on Human Research Ethics* **16(3)**, pp.165-178. <http://doi.org/10.1177/1556264621992214>
15. Darquy S. et al. (2016) 'Patient/family views on data sharing in rare diseases: study in the European LeukoTreat project', *European Journal of Human Genetics* **24**, pp. 338-343. <https://doi.org/10.1038/ejhg.2015.115>
16. De Freitas C. (2021) 'Public and patient involvement in health data governance (DATAGov): protocol of a people-centred, mixed-methods study on data use and sharing for rare diseases care and research', *BMJ Open* **11**. <http://doi.org/10.1136/bmjopen-2020-044289>
17. Despotou G. et al. (2020) 'Evaluation of patient perception towards dynamic health data sharing using blockchain based digital consent with the Dovetail digital consent application: A cross sectional exploratory study', *Digital Health*, <http://doi.org/10.1177/2055207620924949>
18. Elger B. et al. (2010) 'Strategies for health data Exchange for secondary, cross-institutional clinical research', *Computer Methods and Programs in biomedicine* **99(3)**, pp. 230-251. <https://doi.org/10.1016/j.cmpb.2009.12.001>
19. European Commission, College of Europe. (2018) 'Consultation: Transformation Health and Care in the Digital Single Market', *Synopsis Report*. <https://www.doi.org/10.2759/18589>
20. Evans B. (2012), 'Much Ado About Data Ownership', *Harvard Journal of Law & Technology* **25**, pp. 69-130. <http://jolt.law.harvard.edu/articles/pdf/v25/25HarvJLTech69.pdf>
21. Ford E. et al. (2019) 'Our data, our society, our health: a vision for inclusive and transparent health data science in the United Kingdom and beyond', *Learning Health Systems*. <https://doi.org/10.1002/lrh2.10191>
22. Franklin E., Nichols H., House L., Buzaglo J. (2020) 'Cancer Patient Perspectives on Sharing of Medical Records and Mobile Device Data for Research Purposes', *Journal of Patient Experience* **7(6)**, pp.1115-1121. <https://doi.org/10.1177/2374373520923837>
23. Ghafur S. et al. (2020) 'Public perceptions on data sharing: key insights from the UK and the USA', *Digital Health* **2(9)**, pp. 444-446. [https://doi.org/10.1016/S2589-7500\(20\)30161-8](https://doi.org/10.1016/S2589-7500(20)30161-8)
24. Gordon W., Catalini C. (2018) 'Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability', *Computational and Structural Biotechnology Journal* **16**, pp. 224-230. <https://doi.org/10.1016/j.csbj.2018.06.003>
25. Grande D., Mitra N., Shah A., Wan F. (2013) 'Public preferences about secondary uses of electronic health information', *JAMA Internal Medicine* **173(19)**, pp.1798-1806. <http://doi.org/10.1001/jamainternmed.2013.9166>
26. Hafen E. (2019), 'Personal Data Cooperatives – A New Data Governance Framework for Data Donations and Precision Health'. In: Krutzinna J., Floridi L. (eds) *The Ethics of Medical Data Donation*. Philosophical Studies Series, vol 137. Springer, Cham. [https://doi.org/10.1007/978-3-030-04363-6\\_9](https://doi.org/10.1007/978-3-030-04363-6_9)
27. Hassan L., Nenadic G., Tully M. (2021), 'A social media campaign (#datasaveslives) to promote the benefits of using health data for research purposes: mixed methods analysis', *Journal of Medical Internet Research*, **23(2)**. <http://doi.org/10.2196/16348>

28. Hay A. et al. (2017) 'Linkage of clinical trial and administrative data: a survey of cancer patient preferences', *Current oncology* **24(3)**, pp.161-167. <https://doi.org/10.3747/co.24.3400>
29. Heggul N. et al. (2019) 'In response to Ballantyne and Shaefer's 'Consent and the ethical duty to participate in health data research'', *Journal of Medical Ethics* **45**, pp. 351-352. <http://dx.doi.org/10.1136/medethics-2018-105271>
30. Howe N., Giles E., Newbury-Birch D., McColl E. (2018), 'Systematic review of participants' attitudes towards data sharing: a thematic synthesis', *Journal of Health Services Research & Policy* **23(2)**, pp. 123-133. <http://doi.org/10.1177/1355819617751555>
31. Joly Y. et al. (2016) Are data sharing and privacy protection mutually exclusive? *Cell* **167(5)**, pp. 1150-1154. <http://dx.doi.org/10.1016/j.cell.2016.11.004>.
32. Jones K. et al. (2019) 'Public Views on models for accessing genomic and health data for research: mixed methods study', *Journal of Medical Internet Research* **21(8)**, e14384. <https://doi.org/10.2196/14384>
33. Jumbe N. L., Murray, J. C. and Kern, S. (2016) 'Data Sharing and Inductive Learning — Toward Healthy Birth, Growth, and Development', *New England Journal of Medicine* **374(25)**, pp. 2415–2417. <https://doi.org/10.1056/NEJMp1605441>
34. Kalkman S. et al. (2019a) 'Responsible data sharing in international health research: a systematic review of principles and norms', *BMC Medical Ethics* (**20**), article 21. <https://doi.org/10.1186/s12910-019-0359-9>
35. Kalkman S. et al. (2019b) 'Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence', *Journal of Medical Ethics*. <http://doi.org/10.1136/medethics-2019-105651>
36. Karampela M. et al. (2019a) 'Exploring users' willingness to share their health and personal data under the prism of the new GDPR: implications in healthcare', *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) in Germany*, pp. 6509-6512. <http://doi.org/10.1109/EMBC.2019.8856550>
37. Karampela M. et al. (2019b) 'Connected health user willingness to share personal health data: questionnaire study', *Journal of Medical Internet Research* **21(11)**, e14537. <https://doi.org/10.2196/14537>
38. Knaup P. et al. (2018) 'Implementation of a National Framework to Promote Health Data Sharing - The German Medical Informatics Initiative', *IMIA Yearbook of Medical Informatics 2018*, pp. 302-304. <https://d-nb.info/1179024214/34>
39. Knoppers B. M. and Joly Y. (2018) 'Introduction: the why and whither of genomic data sharing', *Human Genetics* **137(8)**, pp. 569–574. <http://doi.org/10.1007/s00439-018-1923-y>
40. Knoppers B. M. and Thorogood, A. M. (2017) 'Ethics and Big Data in Health', *Current Opinion in Systems Biology* **4**, pp. 53–57. <http://doi.org/10.1016/j.coisb.2017.07.001>
41. Kostkova P. et al. (2016) 'Who owns the data? Open data for healthcare', *Frontiers in Public Health* **4**, article 7. <http://dx.doi.org/10.3389/fpubh.2016.00007>
42. Kuntsman A., Miyake, E. and Martin, S. (2019) 'Re-thinking Digital Health: Data, Appisasion and the (im)possibility of 'Opting out'', *Digital Health*. <http://doi.org/10.1177/2055207619880671>
43. Kuntz R. et al. (2019) 'Individual patient-level data sharing for continuous learning: a strategy for trial data sharing', *NAM perspectives*. <https://doi.org/10.31478/201906b>

44. Lamas E. et al. (2015) 'Ethical, Legal and Social Issues related to the health data-warehouses: re-using health data in the research and public health research', *Stud Health Technol Inform* **210**, pp. 719-723. <https://pubmed.ncbi.nlm.nih.gov/25991247/>
45. Lemieux V. L. et al. (2021) 'Having Our "Omic" Cake and Eating It Too?: Evaluating User Response to Using Blockchain Technology for Private and Secure Health Data Management and Sharing', *Frontiers in Blockchain* **3**. <http://doi.org/10.3389/fbloc.2020.558705>
46. Lounsbury O. et al. (2021) 'Opening a "Can of Worms" to Explore the Public's Hopes and Fears About Health Care Data Sharing: Qualitative Study', *Journal of Medical Internet Research* **23(2)**, e22744. <https://doi.org/10.2196/22744>
47. Mählmann L. et al. (2017) 'Big Data for Public Health Policy-Making: Policy Empowerment', *Public Health Genomics* **20(6)**, pp.312–320. <http://doi.org/10.1159/000486587>
48. McCoy, M. S., Joffe, S. and Emanuel, E. J. (2020) 'Sharing Patient Data Without Exploiting Patients', *JAMA*, **323(6)**, pp. 505–506. <http://doi.org/10.1001/jama.2019.22354>
49. McKeown A. et al. (2021) 'Ethical Issues in Consent for the Reuse of Data in Health Data Platforms', *Science and Engineering Ethics* **27**, article 9. <https://doi.org/10.1007/s11948-021-00282-0>
50. Middleton A. et al. (2020) 'Global Public Perceptions of Genomic Data Sharing: What Shapes the Willingness to Donate DNA and Health Data?', *The American Journal of Human Genetics* **107(4)**, pp. 743–752. <http://doi.org/10.1016/j.ajhg.2020.08.023>
51. Mirchev M., Mircheva I., Kerekovska A. (2020) 'The Academic Viewpoint on Patient Data Ownership in the Context of Big Data: Scoping Review', *Journal of Medical Internet Research* **22(8)**, e22214. <https://doi.org/10.2196/22214>
52. Neves A. L. et al. (2019) 'Health Care Professionals' Perspectives on the Secondary Use of Health Records to Improve Quality and Safety of Care in England: Qualitative Study', *Journal of Medical Internet Research* **21(9)**, e14135. <https://doi.org/10.2196/14135>
53. O'Connor Y. et al. (2017) 'Privacy by Design: Informed Consent and Internet of Things for Smart Health', *Procedia Computer Science* **113**, pp.653-658. <https://doi.org/10.1016/j.procs.2017.08.329>
54. Patil S. et al. (2016), 'Public preferences for electronic health data storage, access, and sharing — evidence from a pan-European survey', *Journal of the American Medical Informatics Association* **23 (6)**, pp. 1096-1106. <http://doi.org/10.1093/jamia/ocw012>
55. Petersen C. (2016) 'The Future of Patient Engagement in the Governance of Shared Data', *The Journal of Electronic Health Data and Methods* **4(2)**. <https://www.doi.org/10.13063/2327-9214.1214>
56. Pickard K. et al. (2014) 'Big Desire to share big health data: a shift in consumer attitude toward personal health information', *Research Gate*. <http://doi.org/10.13140/2.1.2107.2960>
57. Porsdam Mann S., Savulescu J. and Sahakian B. J. (2016) 'Facilitating the ethical use of health data for the benefit of society: electronic health records, consent and the duty of easy rescue', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374(2083)**. <https://doi.org/10.1098/rsta.2016.0130>
58. Prey J. E. et al. (2016) 'International perspectives on sharing clinical data with patients', *International Journal of Medical Informatics* **86**, pp. 135–141. <https://www.doi.org/10.1016/j.ijmedinf.2015.11.007>
59. Riggs E. R. et al. (2019) 'Development of a consent resource for genomic data sharing in the clinical setting', *Genetics in Medicine* **21(1)**, pp. 81–88. <https://www.doi.org/10.1038/s41436-018-0017-5>

60. Satinsky E. et al. (2018) 'Mental health service users' perceptions of data sharing and data protection: a qualitative report', *J Innov Health Inform* **25(4)**, pp. 239-242. <http://dx.doi.org/10.14236/jhi.v25i4.1033>
61. Sarkar U, Lyles C. (2020) 'Devil in the details: understanding the effects of providing electronic health record access to patients and families', *BMJ Quality & Safety* (**29**), pp. 965-967. <http://dx.doi.org/10.1136/bmjqs-2020-011185>
62. Seltzer E. (2019) 'Patients' willingness to share digital health and non-health data for research: a cross-sectional study', *BMC medical informatics and decision making* **19** (**157**). <https://doi.org/10.1186/s12911-019-0886-9>
63. Shah N. et al. (2019a) 'Motivations for data sharing—views of research participants from four European countries: A DIRECT study', *European Journal of Human Genetics* **27(5)**, pp. 721–729. <https://www.nature.com/articles/s41431-019-0344-2>
64. Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG. (2016) 'Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study', *J Med Internet Res* **18(4)**. <https://doi.org/10.2196/jmir.5011>
65. Stepanov I. (2020) 'Introducing a property right over data in the EU: the data producer's right – an evaluation', *International Review of Law, Computers & Technology*, **34** (1), pp. 65-86. <https://www.tandfonline.com/doi/full/10.1080/13600869.2019.1631621>
66. Tully M. P. et al. (2018) 'Investigating the Extent to Which Patients Should Control Access to Patient Records for Research: A Deliberative Process Using Citizens' Juries', *Journal of Medical Internet Research* **20(3)**. <https://doi.org/10.2196/jmir.7763>
67. Vayena E., Salathé M., Madoff L., Brownstein J. (2015), Ethical challenged of big data in public health, *PLoS Computational Biology*, **11(2)**. <https://doi.org/10.1371/journal.pcbi.1003904>
68. Vitak J. et al. (2018) 'Privacy Attitudes and Data Valuation Among Fitness Tracker Users', *iConference*, pp. 229–239. [https://www.doi.org/10.1007/978-3-319-78105-1\\_27](https://www.doi.org/10.1007/978-3-319-78105-1_27)
69. Wiggins A., Wilbanks J. (2019) 'The Rise of Citizen Science in Health and Biomedical Research', *The American Journal of Bioethics* **19** (**8**). <https://doi.org/10.1080/15265161.2019.1619859>
70. Williams H. et al. (2015) 'Dynamic Consent: A Possible Solution to Improve Patient Confidence and Trust in How Electronic Patient Records Are Used in Medical Research', *JMIR Medical Informatics* **3(1)**. <https://www.doi.org/10.2196/medinform.3525>
71. Winickoff E., Jamal L., Anderson N.R (2016) 'New modes of engagement for big data research', *Journal of Responsible Innovation* **3(2)**, pp. 169-177. <https://doi.org/10.1080/23299460.2016.1190443>
- Political statements and EU legislation
72. Charter of Fundamental Rights of the European Union 2012/C 326/02  
<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012P%2FTXT>
73. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data  
<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31995L0046>
74. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act), COM/2020/767 final

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767>

75. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

<https://eur-lex.europa.eu/eli/reg/2016/679/oj>

#### Others

76. “Data Access Compliance Office (DACO)”, *International Cancer Genome Consortium's website*, accessed on October 12, 2021. <https://daco.icgc.org/>

77. “Découvrir Google Scholars”, *presentation of the Institut National de la Recherche Agronomique, lastly updated in April 2020*, accessed on July 12, 2021.

[https://ist.inrae.fr/wp-content/uploads/sites/21/2021/01/Infodoc-Google\\_scholar2021.pdf](https://ist.inrae.fr/wp-content/uploads/sites/21/2021/01/Infodoc-Google_scholar2021.pdf)

78. Deuxième enquête Chaire VIPIP-Médiamétrie (2019), “Données personnelles et confiance : évolution des perceptions et des usages post RGPD”, accessed on July 19, 2021.

<https://cvpip.wp.imt.fr/2019/09/26/31-octobre-2019-18e-rencontre-deuxieme-enquete-cvpip-mediаметrie/>

79. “European Health Data Space”, *EU Commission website*, accessed on October 18, 2021.

[https://ec.europa.eu/health/ehealth/dataspace\\_en](https://ec.europa.eu/health/ehealth/dataspace_en)

80. Gaillard B. “Données personnelles, que prévoit l'Union européenne ?”, *Toute l'Europe internet website*, published on December 16, 2020, accessed on July 8, 2021.

<https://www.touteurope.eu/economie-et-social/donnees-personnelles-que-prevoit-l-union-europeenne/>

81. Legislative Train Schedule – A Europe fit for the Digital Age – Data Act”, *European Parliament's website*, accessed on October 28, 2021. <https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-data-act>

82. Netherlands institute for health services research, Health Connect Partners, Royal College of Surgeons in Ireland, Tilburg University, MLC Foundation (2021). “Assessment of the EU Member States' rules on health data in the light of GDPR” *Specific Contract n° SC 2019 70 02 in the context of the Single Framework Contract Chafea/2018/Health/03*.

[https://ec.europa.eu/health/sites/default/files/ehealth/docs/ms\\_rules\\_health-data\\_en.pdf](https://ec.europa.eu/health/sites/default/files/ehealth/docs/ms_rules_health-data_en.pdf)

83. “TEHDAS identifies barriers to data sharing”, *TEHDAS website*, published on July 8th, 2021, accessed on July 29, 2021.

<https://tehdas.eu/news/tehdas-identifies-barriers-to-data-sharing/>

84. TEHDAS, “Why health is a special case for data governance”, published on June 23, 2021.

<https://tehdas.eu/news/tehdas-analysis-health-data-needs-dedicated-eu-regulation/>

85. TEHDAS, “M5.8 : Potential health data governance mechanisms for European Health Data Space” published on September 1, 2021.

<https://tehdas.eu/news/eu-should-rethink-policies-on-health-data-access/>.

## 5 Annex

### List of contributors

Author	Partner
Wannes van Hoof	Sciensano
Alexandre Romenville	Health Data Hub
Rosie Richards	NHS Confederation