Healthy Data

Louise Mathieu
Marlies Saelaert
Wannes Van Hoof

-

Sciensano – Cancer Centre

-

2023

HEALTHY DATA
TABLE OF CONTENTS

HEALTHY DATA

1. What is Health Data ?

Whenever we go to a doctor or a hospital, they collect data about us, our health and our lifestyle, this is recorded and stored in our patient record. It may include our height and weight, whether we smoke, how much we drink, details of any allergies, what aches, pains or infections we've got, and what medications we are taking. It may also include the results of blood tests, images from MRI scans, and any procedures we've had, together with contact information, date of birth, and next of kin information. Other specialists we see, for example dentists, physiotherapists and psychologists will also create records.

Health systems use this information to help provide the best clinical care for us. But health data can also be reused for purposes beyond direct care including to improve health, care and services through research and planning.

Reusing health data has a huge potential to make better use of information to implement the right changes and improve public health and healthcare. Without reusing it, we would not be able to analyse problems and look for solutions. In this sense, data can directly help us as individuals through the care we receive, but at the same time it can also benefit other individuals.

What's more, none of this would be possible without your data, let alone the data of others. Indeed, individual health data can have a real impact once it is shared and compared with that of others..

WHAT ARE THE DIFFERENT TYPES OF HEALTH DATA?

Information created by healthcare professionals

- Electronic health records, which can include results of your blood test analyses, your medical prescriptions, the list of your vaccines.
- National healthcare databases holding items like prescriptions, information about your laboratory tests, and in some countries details about important diseases like cancer.

Information created by patients

- Monitoring illnesses using computer or mobile phone applications.
- Wearable devices such as smartwatches for fitness monitoring and tracking changes relevant to medical conditions.
- How people prevent illness or detect illnesses early, such as screening tests and dietary monitoring.
- Social media posts can be analysed anonymously in aggregate form, for example, to discover how many people are discussing certain side effects of a new treatment.

Wider information that can affect people's health, such as climate and pollution monitoring data

WHAT CAN WE EXPECT FROM THE REUSE AND SHARING OF HEALTH DATA

Watch the video <u>Patient data saves lives: Cancer - YouTube</u>
Reusing health data has a huge potential to make better use of information to implement the right changes and improve public health and healthcare:

- From the patients' perspective, they would not have to undergo similar procedures or repeat their health histories since their healthcare provider would directly access their health information.
- For researchers, reusing health data could also be time-saving and avoid duplicating efforts. Rather than collecting data from research participants, scientists could directly access data already available.
- Policy-makers can also implement more-informed and real-time policy, when it is needed (such as during the COVID-19 pandemic).

Without data we would not be able to analyse problems and look for solutions. Using data for improving health, care and services through research and planning is then necessary to preserve and improve common goods such as public health.

WHO SHOULD BENEFIT FROM IT?

Your data can help you directly, as an individual, to provide you with care, but it can also help other people's lives by benefiting everyone else. In the following video, you can find out how our data can help you and those you love, as well as everyone else.

Watch the video <u>Patient data saves lives: Cancer - YouTube</u>

HOW IS THE VALUE OF DATA CREATED, AND BY WHO?

While all this would not be possible without our data, it would not be possible without the data of others either. Indeed, individual health data has a practical impact only once it is linked and compared with other individuals health data. To understand how valuable health data is created, we need to understand the difference between:

Individual data: A takes treatment X, B takes treatment X etc...

...**and aggregate data**: In this population, 23% of people take treatment X.

Simply, if small amounts of data from many people are linked up and pooled, researchers and doctors can look for patterns in the data, helping them develop new ways of predicting or diagnosing illness, and identify ways to improve clinical care. This would not be possible by analysing the data of one single person. Indeed, it is only once we have compared our individual health data with others that these patterns can appear. Citizens bring their individual data, like a "raw material", and healthcare professionals and researchers will then add substantial value to it, by pooling them together, interpreting them and analysing trends. Therefore, the value of health data is the result of a collaborative and multi-staged process among different actors.

HOW DOES IS AFFECT CITIZENS?

People can sometimes feel that data belongs to them. Our legal framework in the European Union recognizes that health data is highly sensitive and confidential information about us and our lives, which is why it deserves protection, why citizens have rights, and why we might choose to engage in some way in the management of data about us. However, in addition to the fact that the value of our data has a collective dimension, the European regulatory framework on data protection does not recognize that we individually own our health data. The property of health data has been a hot topic of discussion among the academic community. Above all, conceiving data as a property does not legally imply that you can do whatever you want with it.

2. What is health data used for, and by who?

Whenever we go to a GP, a hospital or a pharmacy, information will be collected about us and our medical history. Only healthcare professionals who are directly involved in your care will be able to access your full patient record which includes your medical history. However, some of the information from your record may also be useful for specific purposes beyond your individual care, to improve health, care and services through research and planning. People should be able to find out what's allowed and what's not, and how the data is kept safe.

There can be multiple purposes for using data, such as improving the health system and public policy, understanding the causes and risks of diseases, diagnosis, patient safety, as well as supporting individual care, the development of new treatments and the prevention of diseases.
Only health professionals have access to medical records. Strict controls are imposed on anyone else wishing to access this information for purposes beyond individual care, such as university researchers or private companies.

Some national organisations are responsible for monitoring and deciding which purposes can be pursued.

TYPES OF USE BEYOND DIRECT CARE

Whenever we go to a doctor or a hospital, they collect data about us, our health and our lifestyle. This is recorded and stored in our patient record. National health systems use this information to help provide the best clinical care for us.
Your health data can also be used for purposes beyond direct care including to improve health, care and services through research and planning. This information can be used to help

INDIVIDUELE ZORG

BETERE DIAGNOSES

ZIEKTEN BEGRIJPEN

DATA GEBRUIKEN OM GEZONDHEID EN ZORG TE VERBETEREN

BEHANDELING EN PREVENTIE

BELEID EVALUEREN

GEZONDHEIDSZORG VORMGEVEN

VEILIGHEID VAN DE PATIËNT
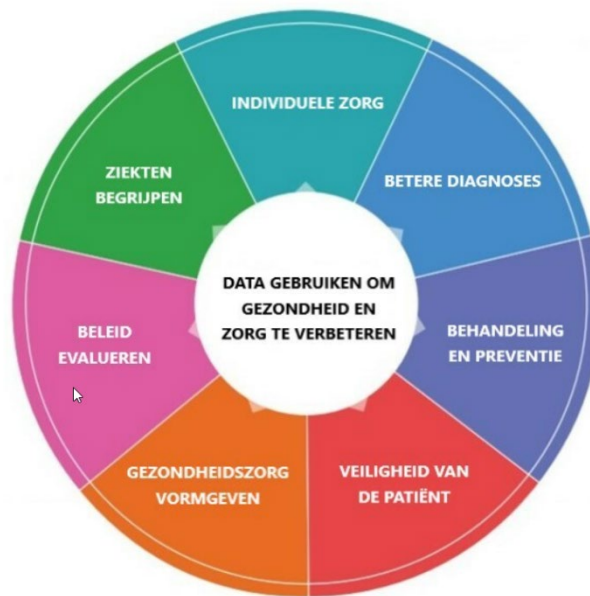
1. Understanding disease

such as why diabetes rates vary between ethnic groups; or finding out about the biological changes involved in Parkinson's disease..

2. Individual care

such as making sure your whole care team, from GP to hospital to care home, has up-to-date information to help give the best and safest care; or helping people to manage their condition, from psychosis to asthma, using smart phone apps

3. Improving diagnosis

Such as demonstrating that bowel cancer screening is effective; or finding new ways to identify early warning signs of dementia

4. Treatment and prevention

Such as testing a treatment for lung disease in the real world; or investigating the benefits of giving statins to men who haven't had a heart attack

5. Patient safety

such as monitoring the safety of whooping cough vaccination during pregnancy; or checking the safety of hip replacements

6. Planning national health services

such as finding out how many people have arthritis to inform healthcare planning; or improving the delivery of kidney dialysis services

7. Evaluating policy

such as comparing cancer survival rates between countries; or showing the impact of a smoking ban on reducing premature births

WHO CAN ACCESS PATIENT DATA?

Your full patient record will only be seen by healthcare professionals who are directly involved in your care or because you provided them with the data.
There are strict controls on how anyone else can access patient information, for purposes beyond your individual care. The purpose must be approved before anyone can use data, and they are only given access to the minimum amount of data necessary. The types of organisations that can use patient data include:

- National health service providers and commissioners: use data to monitor trends and patterns in hospital activity, to assess how care is provided, and to support local service planning.

- University researchers: use data to understand more about the causes of disease, to develop new ways of diagnosing illness or to identify ways to develop new treatments.

- Charities: use data to evaluate services and identify ways to improve care.

- Companies: in certain cases or countries, they use data to provide care and research

How are decisions made about who can access patient data?

Verschillende sleutelorganisaties bewaren op het nationale niveau patiëntengegevens en/of hebben de verantwoordelijkheid om een overzicht te bewaren over de doeleinden waarvoor de gegevens mogen gebruikt worden. Ze werken samen om beslissingen te nemen over hoe de gegevensbescherming kan worden gewaarborgd en om de voorwaarden te bepalen waaronder toegang tot gegevens kan worden verleend.

3. How is health data kept safe?

It is essential that patient data is kept safe and secure, to protect your confidential information. There are four main ways that privacy is protected:
- Removing details that identify a person and taking further steps to de-identify information.
- Using an independent review process to make sure the reason for using patient data is appropriate.
- Ensuring strict legal contracts are in place before data is transferred or accessed.
- Implementing robust IT security.

DATA PROTECTION IS A BALANCING ACT

Do the benefits of using patient data outweigh the risks? Could something go wrong, and what would be the impact? Sharing patient data will never be totally risk-free, but there must be appropriate measures in place to make sure any risk is as low as reasonably possible. Data is de-identified wherever possible. There are audit processes to check who is accessing data, and robust penalties can be issued where data is misused.

CAN I BE IDENTIFIED FROM THE DATA?

Personally identifiable data can only be used if you give your permission or where it is required by law, and even then only with robust safeguards. It cannot be used for insurance or marketing purposes without your consent. Some data will be used to produce statistics that are published monthly by your health care authority, for example hospital emergency waiting times or vaccination coverage. The information can only be openly published if the data is anonymised, so it is not possible to identify any individual.

SPECTRUM OF IDENTIFIABILITY

In practical terms there is a wide spectrum of identifiability. This ranges from fully identifiable personal data, to data that has been through a robust anonymisation process. The bar is very high for data to be considered 'anonymous' under GDPR, which means there are lots of purposes for data use that still count as personal data. The identifiability of data depends both on the features of the dataset and on the environment where it is held and used.
For example, data that is not identifiable on its own may become so if it is combined with other data. Some environments used to store data

Spectrum of identifiability

Personally identifiable — De-personalized — Anonymous

More identifiable ← → Less identifiable

therefore include technical controls on what the data can be linked to and limitations on who can access it. The controls used to protect the data are just as important as the qualities of the data itself.
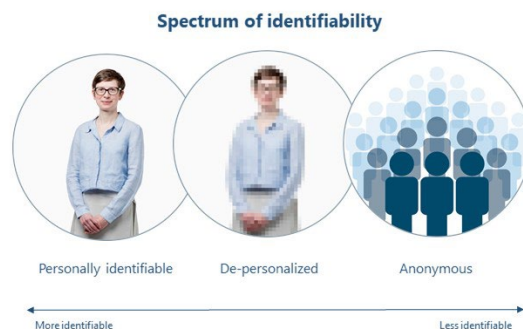
The key issue regarding health data use for purposes beyond your care is the balance that needs to be struck between maximizing the potential benefits and protecting against possible harms. Generally, the utility of health data is highest when few safeguards are imposed, but this also increases the potential risks with regards to privacy protection and the security of the data.
There are two types of health data that are currently treated differently: identifiable and anonymised data.

ANONYMISED DATA

When data is rendered completely anonymous, it is not considered to be a personal data anymore and therefore doesn't fall under the mandate of the General Data Protection Regulation, which only applies to identifiable data. The anonymisation process itself protects individuals from potentially harmful outcomes.

Anonymisation is a continuum: there are different techniques that offer different levels of protection, such as those described below.

Data aggregation

Data aggregation refers to the pooling of data, so that individuals can no longer be identified, such as in the example provided below.

Individual data:
A is fully vaccinated against COVID-19, B is fully vaccinated against COVID-19 etc

...and aggregate data
In this population, 80% of people are fully vaccinated against COVID-19.
.

Data swapping

Data swapping happens where certain characteristics at an individual level are rearranged. This allows researchers to still perform analysis on the entire dataset, but comparisons on an individual level have become meaningless, such as in the following example.

Original dataset

NAME	AGE	CITY	HOSPITAL VISITS	GP VISITS
John	42	London	0	2
Jane	61	Manchester	1	5
Jeremy	21	York	9	0

Same dataset, swapped:

NAME	AGE	CITY	HOSPITAL VISITS	GP VISITS
John	21	Manchester	1	0
Jane	42	York	9	5
Jeremy	61	London	0	2

Small cell risk analysis

A small cell risk analysis is a statistical analysis used to measure the risk of re-identification when only a small group of people is concerned, for example in the case of rare diseases or when many variables are combined.

Advantages and disadvantages related to anonymous data.

- On the one hand, for researchers, using anonymous data is easier because there are little to no restrictions to the use of anonymised data.
  For example, there is no need to ask for participants' consent because no one can be identified. For citizens, anonymous data guarantees that no direct, personal harm - like privacy violation, discrimination or unintentional commercialisation - can occur when data about them is used. However, it is important to note that several publications have shown that sometimes it is possible to reidentify individuals in anonymised datasets. [Click here](#) to see an example.

- On the other hand, anonymous data has less utility. Some research questions cannot be answered without identifying information. For example, it is more difficult to compare across characteristics: if you want to know how a COVID-19 infection affects people with asthma, while correcting for age, high risk jobs and other factors, you need access to all these variables that allow individuals to be identified. For citizens, it is impossible to benefit directly from health data reuse if it is completely anonymised because data quality is reduced and the research participants cannot be recontacted.

IDENTIFIABLE DATA

Identifiable data are all types of data where it is possible to trace back to the individual person behind the data. This includes both:

- **personally identifiable data**, such as name, address, ID number, social security number etc...
- **de-identified or pseudonymised data**, where a person's name and other directly identifying information is removed. It is used very often in health research. There are different techniques of depersonalisation that offer different levels of protection. For example, data can be pseudonymised by replacing some identifiable characteristics or by using a specific encryption (for example, changing names into numbers). A trusted third party can regulate communications between two partners to make sure that neither partner holds all the keys to revert the encryption. However, **it**

may still be possible to re-identify the person if the data is combined with different sources - This would be like adding more pixels to a photograph or joining together different pieces of a puzzle.

All data that is considered to be identifiable is protected by the GDPR. This data can only be used if organisations can show they have a lawful reason for using the data, known as a 'legal basis'. In the EU and the UK, organisations will usually use your data with your 'consent' or without your consent if the use of your data can be considered 'a task in the public interest', or a 'legitimate interest'. If you want to learn more about your rights under the GDPR see *4. How can citizens be involved in data reussage?"*

Informed consent

Informed consent can be used as a legal basis to collect data in a clinical or research context. Asking for consent respects individual autonomy but it can also be burdensome, both for researchers and citizens alike who need to spend time, resources and energy to confirm their preferences every time. Additionally, in some circumstances it may be nearly impossible to obtain consent from some individuals, where their data is needed for a specific project. For example, when data from years ago needs to be accessed to determine trends and evolutions.

A task in the public interest

Informed consent can be used as a legal basis to collect data in a clinical or research context. Asking for consent respects individual autonomy but it can also be burdensome, both for researchers and citizens alike who need to spend time, resources and energy to confirm their preferences every time. Additionally, in some circumstances it may be nearly impossible to obtain consent from some individuals, where their data is needed for a specific project. For example, when data from years ago needs to be accessed to determine trends and evolutions.

Advantages and disadvantages related to identifiable data

- On the one hand, identifiable data offers high utility: it can be used to answer many questions, including when it is necessary to compare certain variables, such as smoking and the effectiveness of a lung cancer treatment. This implies that it has to know who is a smoker and who received a certain treatment. When the use of certain data involves risk to individuals' privacy or other harms, de-identification techniques can be applied to manage these risks (see below).
- On the other hand, it is often a difficult balancing act to decide what the appropriate level of protection is for a certain type of data. Depersonalisation techniques can be time consuming and expensive. For citizens, the use of identifiable data can lead to direct personal benefits, but the risk always remains that individuals can be reidentified when they don't want to be.

OVERSIGHT

Different types of bodies can ensure the uses of health data for purposes beyond individual care are respectful of the regulation and citizens' privacy.

- **Data protection authorities**: data protection authorities are national governance bodies that can oversee structural links that may be created between large databases or decide whether something qualifies as 'a task in the public interest'.

- **Data Protection Officer (DPO)**: some institutions or companies processing personal data under the GDPR need to have a DPO who supports the data processors to make sure that they adhere to the rules described in the GDPR.

- **Ethical Committees**: when new research projects are launched in hospitals or other public institutions, they are submitted to a multidisciplinary ethical committee to decide if the research protocol is ethical and the proposed methods are proportional to the research goals.

- **Data Access Committees**: some databases install data access committees, who decide who gets access to the data and under which conditions.

4. How can citizens be involved in health data reuse?

Did you know that you could have a role in the second life of health data? There are multiple ways through which you could be involved!

BE INFORMED, EXERT YOUR RIGHTS:

One fundamental pillar of UK and European law is to ensure that you are informed on how health data is being processed, by whom, for which purposes, under which safeguards, but also that you know and can exert your rights. Actors using your health information are bound by law to help you exercise your rights and to keep you informed. If they fail to do so, they can be subject to sanctions. You can access, rectify, erase, transfer from one service to another, and object to the use of your related health data.

Your rights

Your right to know about the reuse of your health data
The organisations and people using a part of your health information have an obligation to provide clear and understandable information on what they are doing, including how health data will be used and protected, for which purposes, who will have access to it and for how long they will be using it.
This information provision was enshrined by Law in order to make data users accountable to you: you know what is happening and can thus have a certain control over it.
There are a growing number of training and informative initiatives dedicated to citizens so that they can easily access this information and learn more about what health databases contain and how the use of health data works in practice. Want to see how these initiatives can look like? Visit Understanding Patient Data

Have your say about the reuse of your related health data
This access to information also enables you to exert your rights. Data protection laws guarantee that you can play a role in the reuse of your health data in several ways:
- **You can check whether your health data is included in health databases** which are used for research or other public interest projects. For example, if you have been treated in a hospital, you can check whether the information that you provided to the healthcare professionals that took care of you could be reused. This information is usually indicated on the information note provided to you when you arrived at the hospital, or on its website.

- **You can request to access and, in some cases, transfer a copy of your personal health information** (what is usually called "portability right") from one actor to another. For example, when you want to change your health insurance, you can ask for the transfer of your personal information.
- **You can ask to rectify the data related to you** if you see that there is an error or mistake in your health record.
- **You can express your consent or object to the use of your personal health information**. Under European law, actors using your related health data are committed using a specific legal basis for accessing and using your health data, except in some cases, for example when your personal information is de-identified and used for research projects (more about de-identification: *see 3.How is Health Data kept safe?*). In that situation, you can in most cases express your opposition to it. However, this right cannot apply to the projects necessary for reasons of public interest or which are mandatory according to law.
- **You can ask to erase your personal health information.** Once you express your request, this does not mean that your data will automatically or immediately be erased. For instance, some health records have to be archived (which means that they are not actively used anymore and are kept in a separate database for regulatory compliance or for future uses) during a timeframe enshrined into law and it is impossible to delete them before.

The actors using your health information **are bound by law to help you to exercise your rights.** Some institutions have a dedicated Data Protection Officer to assist you and you can find their contact details on the website or on the information notes provided by healthcare organisations. When there is not a dedicated Data Protection Officer, you can directly contact your healthcare provider or, if you participated in a research project, the person who was in charge of conducting it.

In the event that someone refuses to meet your demand or does not answer your request, you can always **contact the national supervisory authority of your country,** which is the independent organization tasked to ensure the protection and respects of your rights and freedom.

GIVE YOUR OPINION

the use of personal health information is not just a topic for experts in data science or artificial intelligence, it is also a public issue: your opinion and preferences matter and can be taken into account in a number of ways by the actors involved in the use of health data for purposes beyond your individual care. Examples include, polls, open debates, organization of workshops and individual discussions with citizens or their representatives etc.

Citizens' involvement in decision-making

Beyond the rights that you have with regards to what is happening in the second life of your health data, your opinion can also be taken into account in a number of other ways.
- **Consultations:** Actors involved in the reuse of your health data have several tools at their disposal to ask for your opinion and listen to what you have to say. Find out more on these tools and discover some examples of citizens' involvement mechanisms.
- **Polls and questionnaires** can enable you to provide your opinion on specific topics
- **Online and offline open debates**, like this current website for example, can contribute to the development of the public debate around what is happening to health data in its second life and to highlight what matters most to you and the other citizens.

> **Did you know that…** During the research activities that we conducted to prepare this consultation, we discovered the example of a citizens' jury conducted in 2016 in the United Kingdom to discuss with citizens the authorization of the use of health personal information for research purposes. During three days, citizens listened to and questioned various experts before deliberating among themselves and providing their final judgement on the issue. The results were then presented by members of the jury to national and local policymakers. Do you want to know more? The article presenting this project is available here.

- **Workshops or individual interviews :** Some institutions and organisations are also organising workshops or individual interviews among members of the public and representatives from the civil society or patients' associations. These formats can enable an in-depth discussion of specific topics or the co-development and identification of the rules that should be respected by actors wishing to use personal health information.

> **Did you know that…** a research team from Canada developed an application to be used by citizens to manage, control and share health personal information. They organized focus groups with 26 people to test their solution and better understand how such health data management tools could be perceived. You want to know more? the results of their research is available here.

- **Representation in Ethics, Advisory and Data Access Committees :** To ensure that patients' and citizens' interests are being respected, organisations involved in the reuse of your health data can also include citizens and/or their representatives in decision making bodies involved in the regulation of access and use of personal health information.

> **Did you know that…** the International Cancer Genome Consortium has an independent committee that monitors and advises requests from scientists to access health data. It includes experts, scientists and potential data users, but also citizens who have had cancer, and independent lay persons. Its role is to ensure that the potentially identifiable data under the responsibility of the Consortium is only used for public health objectives and by qualified researchers

PARTICIPATE IN RESEARCH PROJECTS

individual citizens, patients' associations and civil society organizations can also be involved in or even initiate research projects to ensure that efforts are being made to advance on topics and issues which matter to them while respecting their values

Citizen's participations in research projects

Research projects can involve and even be initiated by citizens, patients' associations or civil society organizations to ensure that efforts are being made to advance topics and issues which matter to them, while respecting their values. Be it through providing information to researchers, participating in the

collection and analysis of health information, or initiating and co-constructing the protocols of a research project, there are many ways in which you can be involved.

5. Wat are Health Data Spaces?

Each time we go to the hospital, to our general practitioner or to the pharmacy, we leave personal health information behind. This data provides information about us and the state of our health. This data could be a number of things, including a medical prescription or imaging. This data is used and shared between the healthcare professionals who take care of us, to provide us with better care and to ensure the best possible treatment.

But **our data can also be valuable, and even indispensable for research and public health**. Data sharing underpins life-saving research, decision-making and managing health threats such as COVID-19. The more health data is available and accessible for analysis, the better researchers and decision-makers can understand what is happening and develop tailored solutions,

It can still be difficult to share health data, even in emergencies. This is because health stakeholders such as hospitals, general practitioners or researchers collect data in many different formats which are incompatible with one another.

Health data spaces bring together these different types of health data *(see "1.What is Health Data?)* to improve access to and the use of health data by providing common rules which facilitate better data sharing. Health information can be used by various actors for purposes beyond individual care, including research, prevention, diagnosis, and the development of innovations such as new treatments and health services.

WHAT IS THE EUROPEAN HEALTH DATA SPACE?

As the COVID-19 crisis highlighted, health threats know no borders. The European Health Data Space project was conceived by the European Commission and 25 European countries to address the common health issues and challenges we are all facing.

Sharing health data between European countries could help them to address these challenges. The more health data is available and accessible, the better researchers and decision-makers can understand what is happening and develop better and tailored solutions, such as better treatments or better management of health crises.

A common European Health Data Space will promote a better exchange and access to different types of health data, not only for health research and health policy making purposes but also to support patients' care.

6. Case 1 : Healthy data during the COVID-19 pandemic

The COVID-19 pandemic has dominated our lives for more than two years. The battle against the coronavirus has forced us to make changes to the way we work, have fun, educate our children and interact with each other. The justification for these changes is often found in 'the numbers'. But where does this data come from? What is it used for exactly? Which safeguards are in place? And how are different stakeholders and citizens involved in this health data reuse? Individual, societal and stakeholder level choices influence these COVID-19 numbers.
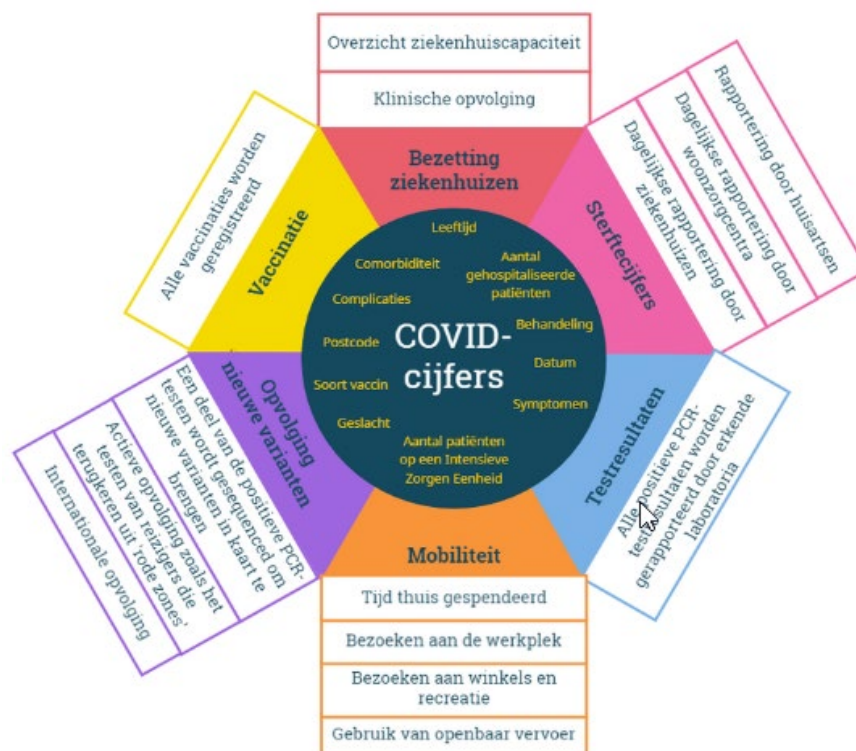
Surge capacity is the ability of a health system to manage a sudden and unexpected influx of patients in a disaster or emergency situation. Creating surge capacity involves a comprehensive approach linking the four S's of surge capacity: space (or structure), staff, supplies and systems. (World Health Organization. Strengthening the health systems response to COVID-19. Creating surge capacity for acute and intensive care. Copenhagen: WHO Regional Office for Europe; 2020)

When a person receives a positive PCR test result, they may need to isolate. They may receive treatment and they can get a proof of past infection. While people get tested to be diagnosed, this data continues to serve other purposes long after the virus has run its course. The second life of the data has important societal implications and allows us to develop policies based on evidence.

It is very important to establish links between different sources of data about COVID-19 to be able to inform policies, develop treatments and prevent the spread of the virus as much as possible. Since the beginning of the pandemic, more than 420.000.000 cases have been confirmed worldwide. To understand each of these cases, one needs to know if the person was vaccinated (when, how often, which vaccine, ...), hospitalized (duration, type of treatment, in the ICU or not, ...), returning from abroad, in contact with other people, infected with which variant of the virus, etc. The links between these different types of data need to be made on an individual level. This means that sensitive, personal health information is being shared. Learn more about the safety of health data in "*3. How Health data is Kept Safe*"

INDIVIDUAL CHOICES

Individual choices are always the starting point of any health data story: only when a person decides to interact with the healthcare system, will there be any data created. However, PCR test results are anonymously and automatically shared, leading to a second life for the data beyond the scope of the individual from whom they originate. This begs the question: should the individual behind the data be directly involved in the reuse of this data? As it stands everyone's' data is used to develop better vaccines, including data from people who do not support vaccination. Data from individuals who are against lockdowns will be used to decide whether to enact them or not, etc. This is the case because everyone's data is grouped together and reused by design, with strict safeguards in place. This ensures robust, timely and complete datasets. However, there are different ways in which citizens could be involved more actively. *See "4. How can citizens be engaged"*

SOCIETAL CHOICES

If individual choices do not impact how COVID-19 data is governed as much, what does the framework for health data reuse look like? In short: personal data is protected by a specific legislation, the General Data Protection Regulation, links between datasets containing personal data are closely managed and under strict review (by Data Protection Authorities, housing a multidisciplinary commission) and anonymised data can be freely used and shared because they cannot be traced back to an individual.
For every type of data, every different purpose and all new links, several considerations need to be made:
- What can the data be used for? Every health data reuse needs to serve a specific purpose that has to be clearly defined.
- Who can access the data and under which conditions? Open access will promote more collaboration, but might require strict safeguards that limit the usefulness of certain datasets.
- Which variables are used: the use of more variables (e.g. age, gender, location, disease, treatment, ...) increases the amount of sensitive information and therefore implies more risks and the requirement of more safeguards. However, the more variables are included, the more useful a dataset becomes.
- How will the data be stored? New standards are being developed to increase findability, accessibility, interoperability and reusability (FAIR-principles). Read more about how health data infrastructure can help in the battle against COVID-19 here

STAKEHOLDERS' CHOISES

All actors in the health data reuse environment make decisions within the framework that is developed at the societal level. They decide which purposes to pursue, which safeguards to specifically employ, which collaborations to establish, etc. For example, the WHO collects data about COVID-19 worldwide and makes it, in an aggregated way, accessible to anyone under an open data licence here. Some public health institutes describe in great detail which data they are using and why (e.g. see the FAQs from Sciensano).

7. Case 2 : We are Health - What if your health data never left your side?

## WHAT THE SITUATION LOOKS LIKE NOW ?

Personal health data is generated every time we interact with the healthcare system and gathered in different places: a medical file, hospital records, health insurance databases, public health databases, universities, etc. This data can later be reused (i.e. for purposes beyond your direct individual care), linked and shared under strict rules and conditions. *See "3. How is health data kept safe?"*

A DIFFERENT PERSPECTIVE

Personal health data could be stored in one place: by the citizens themselves. Such a platform would be giving citizens ownership of their own data and actively involving them in decisions about the data. This gives citizens the guarantee that they can share their data in a safe way as a leverage for both personal wellbeing and the common good. The We Are Health project aims to develop such a platform.

HOW COULD THIS BE REALISED?

Solid technology can provide citizens with personal data vaults, meaning that instead of interacting with servers which store your data externally, your data always stays with you, in your personal pod (for example an app on your phone). Citizens can control who and which applications get access to data in their pod. Data are stored in a standardised and interoperable format, allowing different applications to use the same data. This means that applications always need to come look for your data in your pod, rather than being able to store it on their own servers. The separation of data from applications is a key concept in Solid, challenging the current situation in which organisations collect personal data and keep them under their control for their own use. In short**: your data would no longer be stored in different places, but different users would come looking for the data in your pod, which is under your control.**
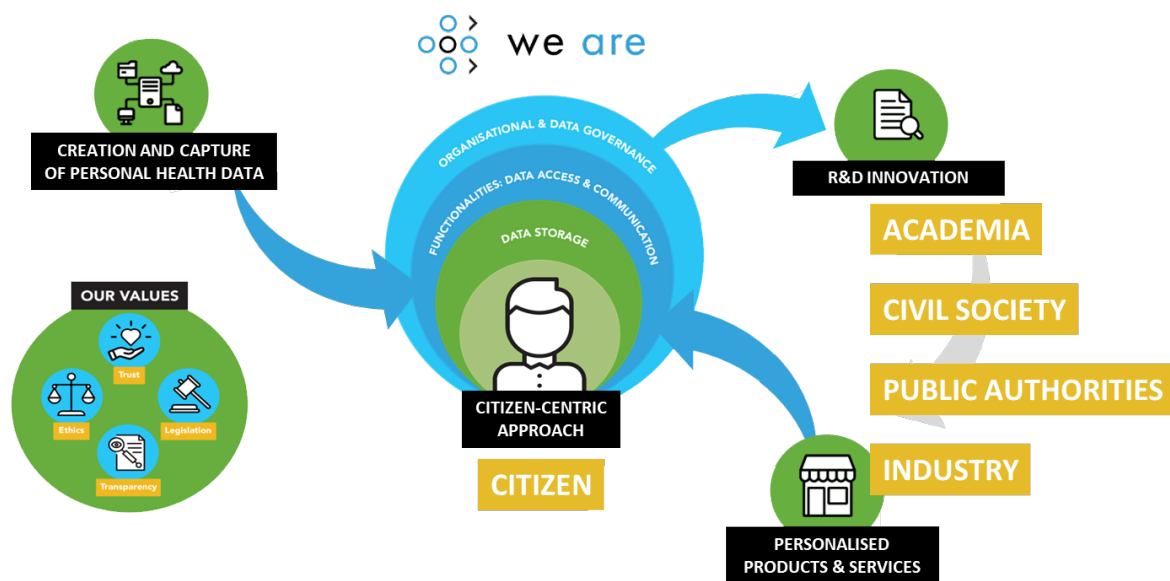
WHAT ARE THE ADVANTAGES OF THIS SYSTEM?

- the citizen is in the driving seat
- the citizen can manage his own data
- there is clarity at all times about what happens to personal data
- the citizen chooses which data (whether or not anonymised) may be used and by whom
- Institutions and/or companies are given new opportunities for research and development
- there will be collaboration with existing platforms for data sharing
- ethics, sustainability, transparency and innovation go hand in hand and are at the service of society

WHAT ARE THE DISADVANTAGES OF THIS SYSTEM?

- it may become harder to obtain data from representative samples of the population
- some research may receive more or less support from citizens, regardless of their importance, risks or merit - for example because of corporate, media or interpersonal hype.citizens would need a lot of data awareness and education
- it may prove difficult to scale this technology

This approach may prove very advantageous for those citizens who want to share their data beyond the current standards of practice. In that case, this kind of platform offers a safe, citizen controlled way forward to more data sharing in a trusted environment, without disrupting current frameworks for health data reuse that improve healthcare and advance scientific knowledge.

8. School participation – Teacher toolkit

Teacher toolkit : Pedagogical Toolbox Healthy Data | sciensano.be
Informative brochure : Informatiebrochure over gezondheidsdata: "Zorg voor je data" | sciensano.be (NL/FR)