

Health Data Basics

This guide uses a pretend example about a visit to a doctor's office to explore how health data can help us understand people's health and improve healthcare.

What are health data?

Health data are tiny pieces of information or facts about your health and well-being. Health data can be about people and/or healthcare services. Health data includes details like the results of medical tests, medicines you've taken, and visits to healthcare providers.

How are health data collected?

Health data are collected in many ways and for specific reasons. Health data are collected through health care interactions, like a visit to a hospital, pharmacy or doctor's office. Health data can also be collected through research, like a study to test how well a drug works.

Who uses health data and why?

Health data can be used in many different ways. For example, by:

- patients, to understand their treatment options.
- healthcare providers, to track treatments that a patient receives and whether they get better.
- researchers, to better understand an illness or develop new treatments.
- people who deliver healthcare, to see how they can improve a service.
- governments, to make decisions about covering the cost of treatments.



Meet Jane...

Jane has bloodwork done to check on her blood sugar. Jane visits her doctor to review the results. The doctor prescribes a new medication (pill) for her diabetes. Jane picks up the pills at her local pharmacy.

Different kinds of data are collected from these visits. For example, the doctor's office uses a billing code (a number that indicates the type of service) and Jane's health card number to get paid for their services. The health card number is linked to some information about Jane like her age and sex. The pharmacy also has a record of Jane's prescription.

These data allow healthcare providers, like Jane's doctor and pharmacist, to see what pills she is taking and check that she receives the right dose.

How can we learn from health data?

Health data can be used to benefit the public. For example, health data can help to answer questions about whether a health service is reaching the people who need it. Often these questions can be answered with health data that already exists.

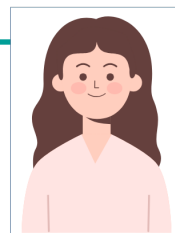
Take the data from Jane's visit to her doctor. Researchers can use these types of data to find out how many people used these services and how often. They can look for patterns, for example based on age. They can see how many people, like Jane, visited their doctors and received the same medication.

This is called **secondary data use** because data collected for one purpose (billing for a service or providing a prescription) is being used for a different reason (understanding how often a service is being used). To protect people's privacy, these health data are de-identified before researchers access them.

What are de-identified data?

When collected, health data includes information that identifies you, like your name and health card number. This is for good reason. In our example, the doctor needs to know who Jane is to care for her.

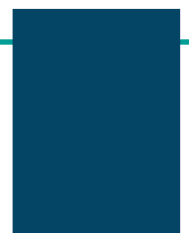
However, researchers do not need to know who Jane is. So, identifiable information, like her name and health card number, is removed and replaced with a random number before researchers see it. These data are called **de-identified data**. This process retains valuable information while protecting privacy.



Fully identifiable



De-identified



Anonymous

Where are health data stored?

Health data are stored in different places, for example a patient's health record. Some health data are stored in secure systems that support secondary data use. These systems can have different names, for example **data repositories, data centres or data platforms**. These systems store large collections of data securely. Researchers are only given access after they apply, explain what data they need to use and how, and have their application approved.

What safeguards are in place?

There are many **safeguards** in place to protect people's privacy and reduce the risk of health data being accessed or used inappropriately. Some examples of safeguards are:

- De-identified data (information that can identify people is removed while retaining valuable information for research).
- Secure facilities that control who and what data is accessed.
- Committees that review and approve how and what data is used.
- Laws and policies about how information can be used.
- Data privacy and security training for researchers before they can access data.

With appropriate safeguards, health data can be used to improve health and healthcare services while protecting privacy.