[Narrator] When you hear the word surveillance, what does it make you think of? Spies? Video surveillance cameras in stores and on buildings? Spies and surveillance cameras are constantly monitoring to collect information that might be useful in some way. The information collected is used to detect problems and help inform actions.

Public health surveillance is similar but uses different tools and strategies to collect information. It starts with experts, like epidemiologists, who constantly monitor what diseases are occurring in the community. They collect and analyze data to identify problems and take action. They review disease data from sources such as laboratories, doctors' offices, and hospitals. When you hear about the number of COVID-19 cases in a state or an outbreak of Ebola in a country, those reports come from public health surveillance.

Like a dog that barks in the night, surveillance systems alert public health experts when something unusual might be going on in a community. Sometimes, a dog barks because someone is outside the door. Sometimes, a dog barks at the wind or nothing at all. The only way to find out is to investigate. Similarly, when a surveillance system detects an increase in cases, public health experts must check the data to take a closer look — it might indicate a real outbreak is occurring, or it could be a false alarm.In setting up a surveillance system, it is important to first determine its purpose and goals. One main reason to do surveillance is to let health experts know when they need to take action. Surveillance data provides answers to questions like "do we have a problem?" and "if so, how big is the problem?"

After taking public health actions to address the problem, surveillance data can also answer, "are the steps we're taking working?" So, the data must be accurate and timely.

When public health experts have identified the purpose and goals of their surveillance system, they then must decide which diseases to include, what data to collect, and how to collect it.

In deciding which diseases to include in their surveillance systems, public health experts in each state develop a list of diseases that health facilities such as hospitals should report to the health department if cases are detected. These are called "reportable diseases" and the cases can usually be confirmed by laboratory testing.

Health departments collect data to monitor these diseases because they can cause serious illness, or can spread easily from person to person, or can cause deadly pandemics. Examples of reportable diseases are rabies, measles, and COVID-19.

Determining which data to collect starts with addressing **case definitions**. Case definitions are used in both surveillance and in outbreak investigations. A case definition is developed for each disease that health experts want to monitor. It provides a set of criteria to define when an instance of disease should be counted as a case. Some case definitions used in surveillance systems are very specific and require laboratory confirmation to ensure that the person has that disease and not something similar. Other case definitions may be written more generally,

requiring only that the illness matches a description of common signs or symptoms.

The case definitions for many reportable diseases are standardized so that the data collected in different places can be compared.

Public health experts collect information about people and their disease including the patient's name, address, phone number, age, sex, gender, race and ethnicity, and occupation. These personal characteristics collected about a person with a case of disease are called demographic information, sometimes "demographics" for short.

Demographic information is generally followed by clinical information such as signs and symptoms, the date the illness began, and the name of the diseases, if known. Signs of a disease are the objective observations noted by a health care provider, such as "jaundice", "rash", or "diarrhea"; and symptoms are subjective reports by the patient, such as "pain", "nausea", or "fatigue".

Information on laboratory testing generally follows clinical findings. This includes samples collected, collection dates, and test results.

Other information may be collected including possible places the person could have been exposed, such as a childcare center, or noting behaviors that could have contributed to a person becoming infected, such as injecting drugs or not washing their hands after being around animals.

Finally, public health experts must consider how best to collect the data. To make this decision, they consider the amount of resources available, like time, funding, staffing, and equipment; the type and possible sources of data; and how and when the data will be used.

The different ways of collecting surveillance data can be generally categorized as passive, syndromic, or active.

Let's begin with an example of a passive surveillance system. Typically, if a doctor diagnoses measles in a child, the doctor will enter her report into an electronic health record system. A series of automatic events will identify this case of measles as a **reportable disease** and automatically send the report to the public health department.

The health department may follow up by calling the patient or their parent, if a minor, to ask additional questions concerning possible exposure and follow-up.

This is a **passive surveillance system** because the health department waits passively for the reports to come in.

Depending on the disease and situation, health departments may also notify other public health agencies, like the U.S. Centers for Disease Control and Prevention or CDC.

This information is added to CDC's National Notifiable Diseases Surveillance System which tracks about 120 diseases across the nation. These include infectious diseases such as Zika, foodborne disease caused by *E. coli*, and noninfectious conditions, such as high levels of lead in a young person's blood. The National Notifiable Diseases Surveillance System helps all levels of public health - local, state, federal - to share health information to monitor, control, and prevent spread of disease.

It can take days or even weeks after a person first feels ill for a doctor to determine what disease that patient has, for lab tests to confirm the diagnosis, and for the report to be sent to the local health department using a passive surveillance system. During this time outbreaks may go undetected and timely intervention may be delayed. A newer type of surveillance, **syndromic surveillance**, was developed to address these challenges.

Syndromic surveillance relies on computer algorithms to review information automatically reported from every health care visit to find patterns of concern. For example, a syndromic surveillance system may identify an increase in sets of signs and symptoms that are linked and consistently associated with a disease, such as fever plus widespread rash. A group of symptoms which consistently occur together are referred to as a syndrome, which is where this system got its name.

An unexpected increase in the number of people experiencing illness could suggest an outbreak that should be investigated. Although further investigation and time may be needed to determine the exact cause of the illness, the power of syndromic surveillance is that it can alert public health experts several days before a problem is detected through passive surveillance. Syndromic surveillance is also useful to help monitor and describe how an event of public health concern is changing over time.

While passive and syndromic surveillance systems can sound the alarm about increases in disease, they are subject to significant under-reporting and may give a false impression of where the problems really lie.

For instance, syndromic surveillance can show there is an increase in the number of people being treated in an emergency room with nausea and vomiting, but it only picks up those people who went to emergency rooms.

To get a more accurate picture of what is going on, public health experts might use an **active surveillance system**. With this system public health experts actively seek out more information by contacting laboratories, hospitals, healthcare providers and other patient care settings to review medical records and identify cases that may have gone unrecognized and unreported.

Public health experts may even set up community testing sites to identify additional cases. These active surveillance systems are typically short-term and are used in outbreak responses. Active surveillance systems give a more accurate picture of the size of the problem and help identify individuals who may need to be treated, isolated or quarantined in order to prevent

additional infections.

After public health experts have collected passive, syndromic, or active surveillance data, they must analyze and interpret the data and decide what actions, if any, are needed. The standard approach for analyzing surveillance data is to summarize the data by **who, where**, and **when**, also referred to as **person, place**, and **time**.

Analyzing data on **person** requires looking at demographic information, such as age, sex, gender, race, ethnicity, and sometimes other characteristics like socioeconomic status, access to medical care, and activities, such as occupation and hobbies.

Analyzing data on **place** involves looking at where the patients live, work, or go to school. Are there clusters in specific areas?

Analyzing data related to **time** involves looking at the number of new cases by day, week or month of onset and determining whether the number of cases is going up, staying steady, or going down and how it compares with other years.

Together, these analyses describe the pattern of disease in a community, help experts decide whether action is needed, and, if so, on whom, where, and when to focus.

Fortunately, Bella, a biostatistician, works at a health department and has all the specialized skills to handle the task. In her work, she uses math to analyze and summarize the collected data. To visualize the data, she creates reports with graphs that show disease patterns or “trends”, area maps that show cases geographically, and bar graphs that show cases by age and sex. Her reports help other public health experts assess disease trends, draw conclusions about specific health problems and actions needed, and develop strategies to control and prevent public health problems.

As a quick recap, we can now answer the question “**Where do public health data come from?**”. Public health data come from surveillance. Surveillance is the ongoing, systemic collection, analysis, and interpretation of data about health conditions in a specific population. Case definitions are used to identify individuals who should be counted as having a case of a certain disease or health condition.

The three main types of public health surveillance include passive, syndromic, and active.

In passive surveillance, the health department relies on healthcare providers or laboratories to report cases.

Syndromic surveillance looks at syndromes, or groups of signs and symptoms, and the data are automatically collected.

Active surveillance is usually short-term and requires public health experts to actively search for

cases.

Information from surveillance helps public health experts describe disease patterns with respect to person, place, and time. That can be used to create strategies to control the disease and prevent future health problems.