Descriptive Epidemiology

Lecturer: Dr. Brandon Guthrie

In this lecture, I will discuss descriptive epidemiology and the types of data collected when conducting a descriptive epidemiologic study.

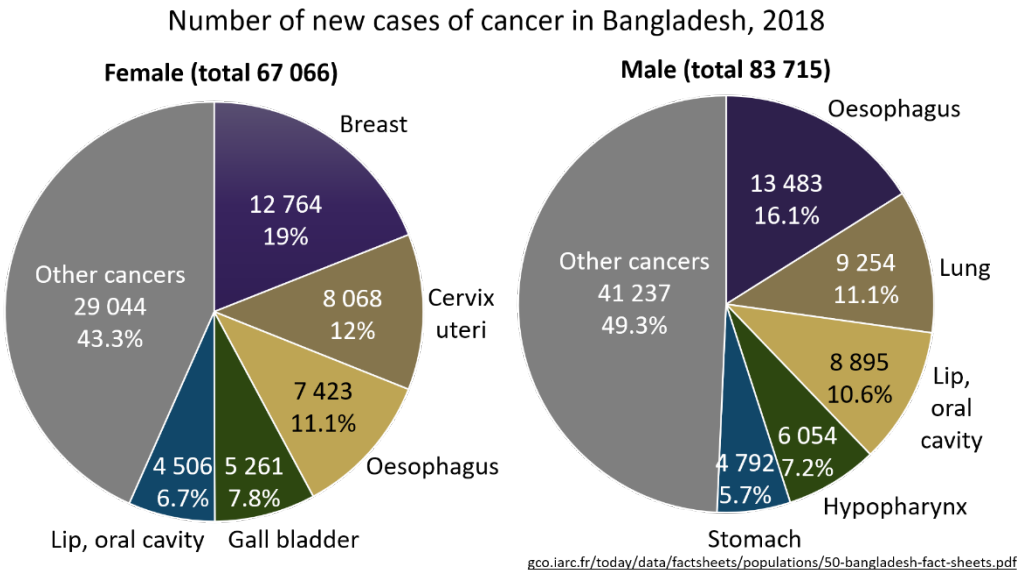
Descriptive epidemiology describes a disease in a population based on person, place, and time. These characteristics of the population by person, place, and time provide important clues about the underlying cause of disease or the source of disease in cases of outbreaks. Based on the information that is available from the descriptive data, it helps to generate hypotheses about a disease. Descriptive epidemiology answers the questions who is getting the disease, what is causing the disease, when is the disease occurring, and where is the disease occurring?

Person

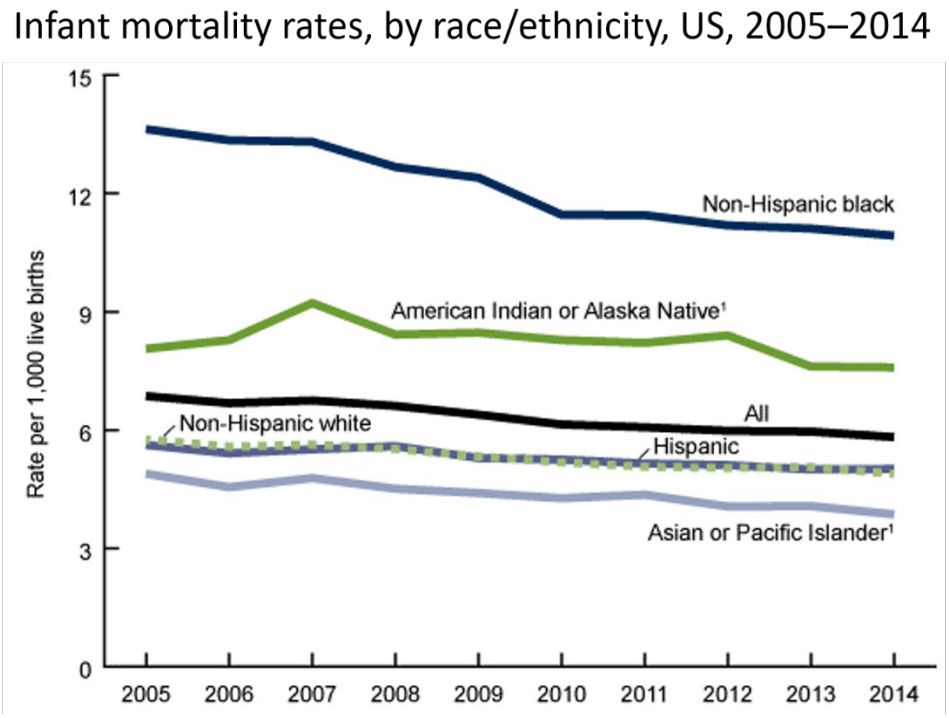
Let's start by looking at person data. Descriptive statistics around people help us identify characteristics of people with disease. Types of person data include age, gender, race, ethnicity, marital status, income, job data, occupation, body weight, travel, religious practices, and many other characteristics. It helps to identify any difference between the diseased people and people who are not diseased with respect to their personal characteristics.

Let's look at some examples.

In these graphs, we see on the left the number of new cases of cancer for all people in Bangladesh in 2018 for females. On the right is the number of new cases for males. We can see quite a lot from these data. We can see that breast cancer has the leading cancer among females followed by cervical cancer. Among males, esophageal cancer was the leading cancer, followed by lung cancer.



This graph shows a difference in infant mortality rate in the US by race or ethnicity. So, we see the time along the y-axis and the infant mortality on the x-axis. There's clearly a higher infant mortality per thousand live births among black individuals compared to other races in the US from 2005 until 2014.

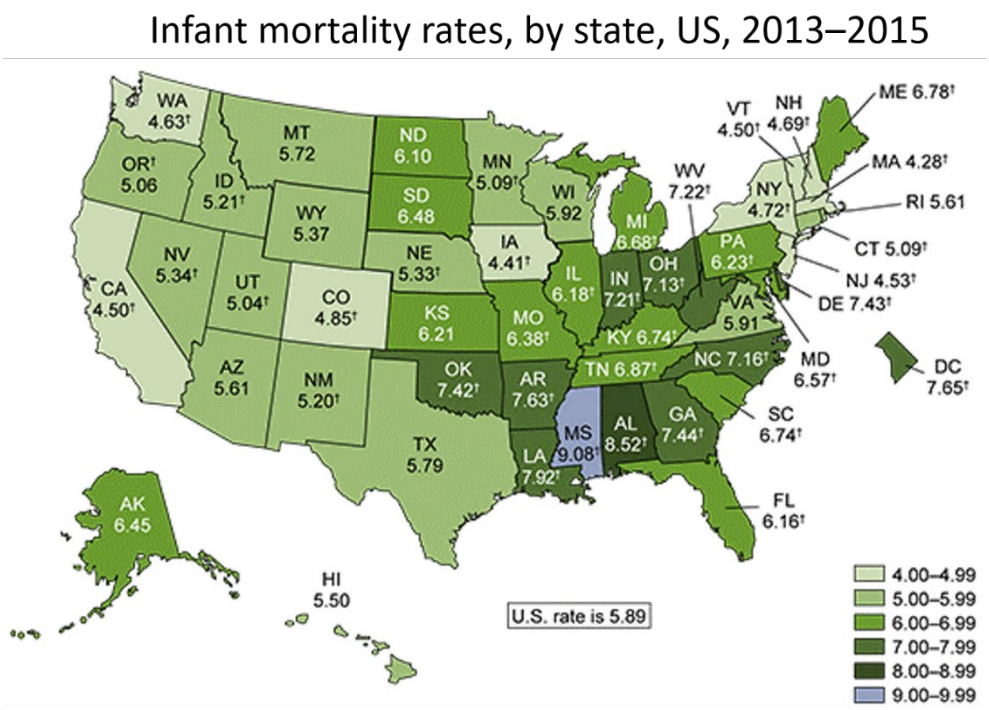


Place

Now, I'd like to talk about place. Place data are equally important in demonstrating the relationship between physical features and disease. Place data provide information on the

geographic extent of the disease. Physical features include altitude, sunlight, temperature, pollution levels, and may also consider proximity to certain environmental features, such as natural resources, hazards, toxins, that may be in a habitat or have a specific pathogen that we might be interested in investigating.

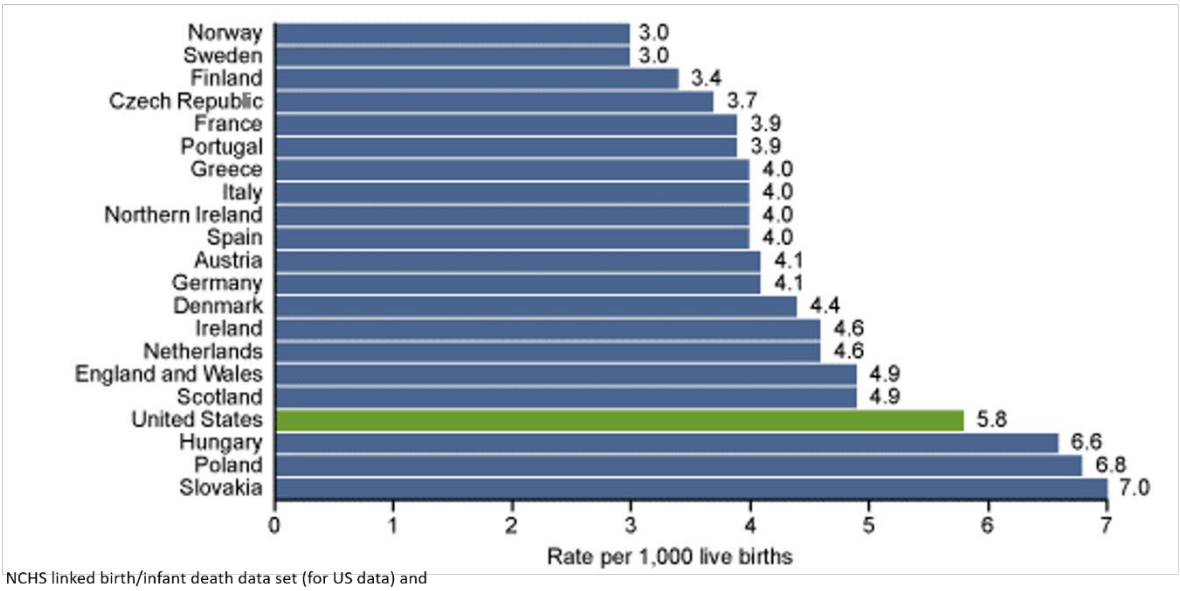
Here are some examples of place data and how we might describe them in a figure. This figure shows variation in the infant mortality rate by state in the US in 2013 to 2015. Mississippi has the highest rate as shown in the purple shade at 9.1 per 1000 live births. Massachusetts has the lowest infant mortality rate at 4.3 per 1000 live births.



https://www.cdc.gov/nchs/products/databriefs/db295.htm

This figure shows the international variation in infant mortality, so we can see quite clearly here with countries along the y-axis that in Norway, infant deaths are lower than in Slovakia, which has the highest infant death rate. These data are powerful, and they communicate a message even without any statistical test, just presenting them a descriptive analysis.

Infant mortality rates Europe and US, 2004

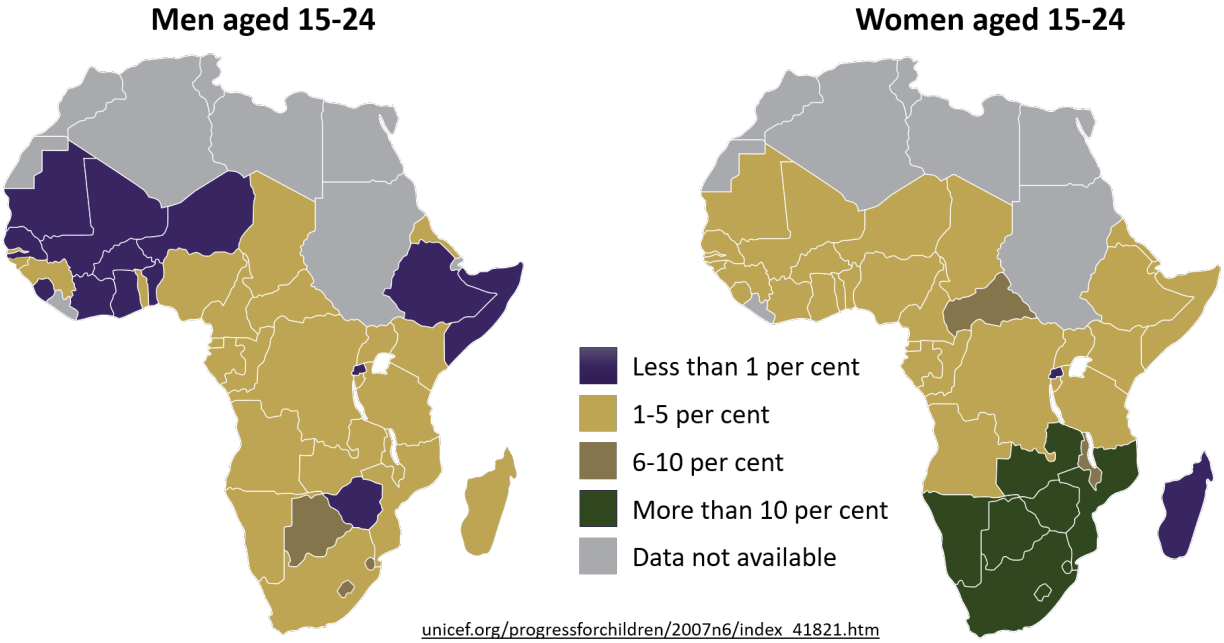


NCHS linked birth/infant death data set (for US data) and *European Perinatal Health Report* (for European data)

https://www.cdc.gov/nchs/products/databriefs/db23.html

Here's another example of place data showing HIV prevalence in men and women aged 15 to 24 in Sub-Saharan Africa. It's clear looking at this map that the epidemic varies across the continent.

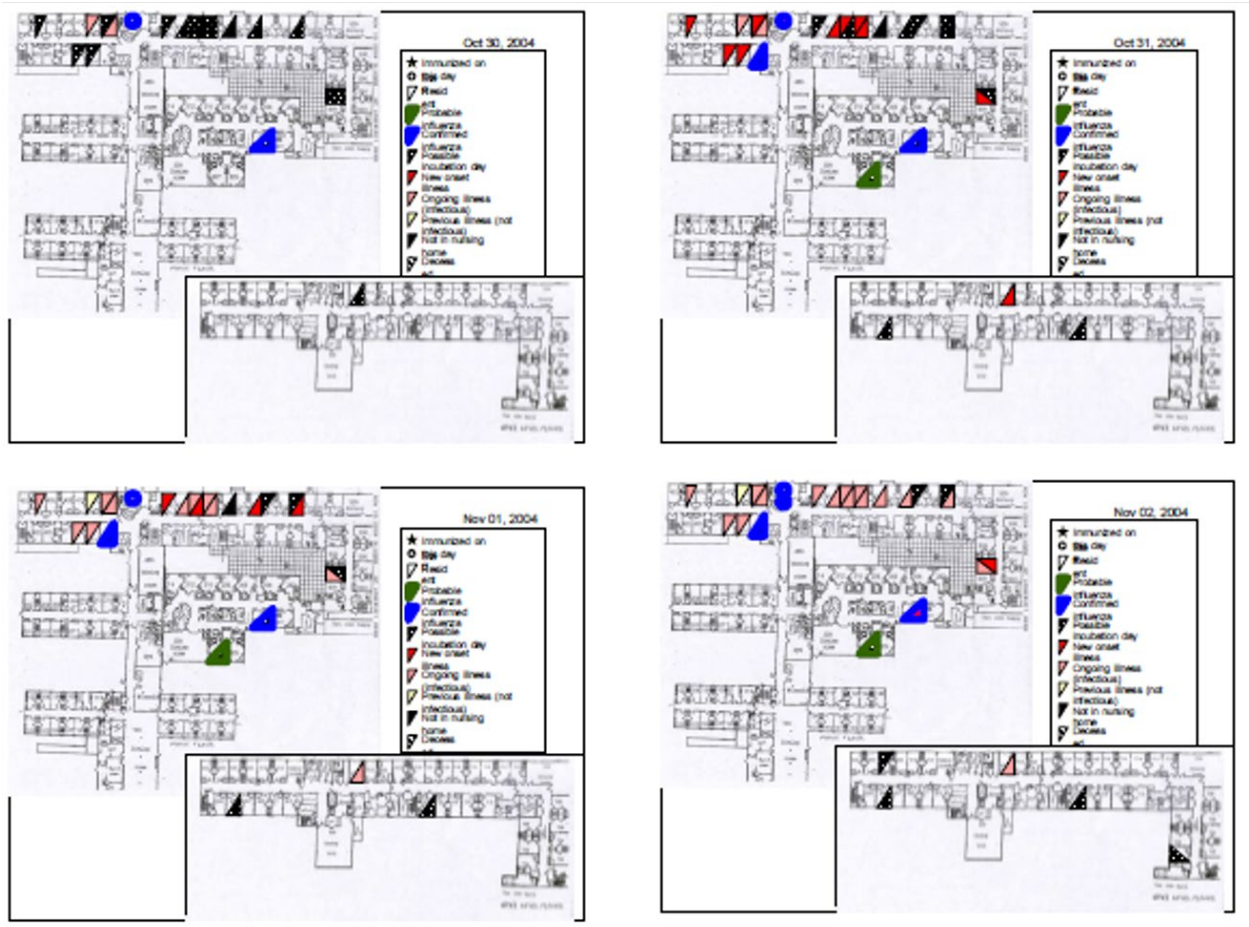
HIV Prevalence in men and women aged 15-24 in sub-Saharan Africa



unicef.org/progressforchildren/2007n6/index_41821.htm

This is another example of place data, but it is not within a country or a continent. The same type of mapping can take place in a physical building such as a nursing home. We can see influenza cases in a nursing home map to determine whether they are concentrated. Is there patient spread? Are there some features of this outbreak that can be used to help us understand the nature of this disease?

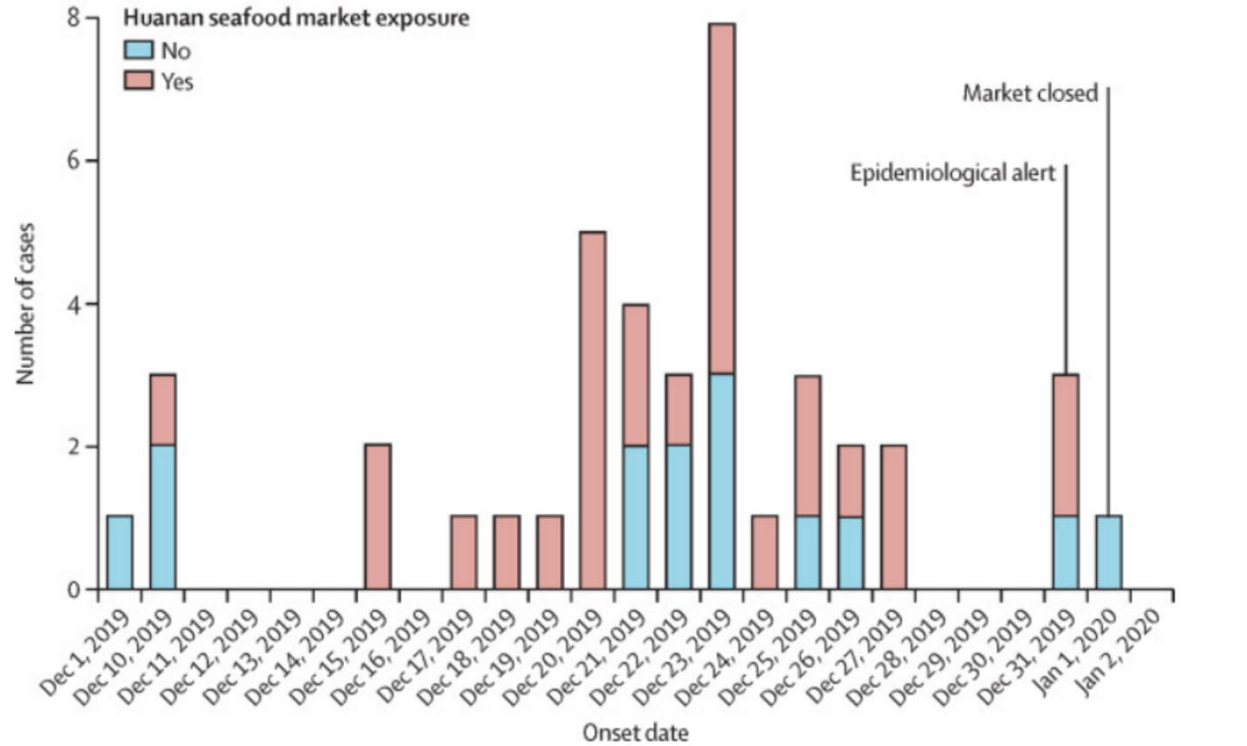
Mapping influenza cases in a nursing home



Time

Now let's talk about time. When we talk about time data, we need to consider different ways, such as calendar time, age, and time, since an event. We also need to look at trends or long-term variation as well as short-term seasonal variations. When we are considering time data, this type of graph is useful to help us identify the cause early on in the outbreak. In this example, the graph shows the date of the illness onset among lab confirmed COVID-19 cases. You can see that the symptom onset date of the first patient identified was December 1, 2019, and there are a cluster of cases around the third week of December.

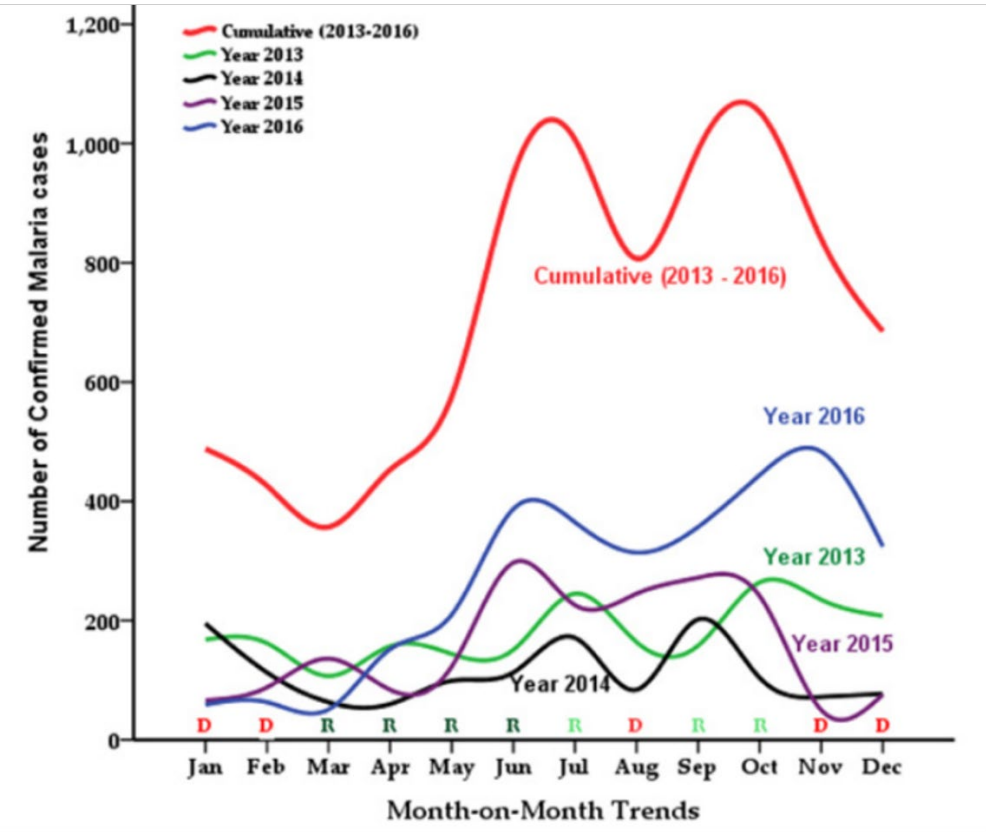
COVID-19 and date of illness onset



thelancet.com/journals/lancet/article/PIIS0140-6736(20)30183-5/fulltext

In this example, we see a month-on-month trend of malaria cases for 2013, 14, 15, and 16. The x axis represents the month and shows it as a dry season (indicated by a D), major rainy season (indicated by R), or minor rainy season (indicated by another R). The general epidemiologic pattern shows a rise in the infection burden each year from June through October and then a decline in the infection burden onward.

Malaria trends in Ghana



benthamopen.com/FULLTEXT/TOMICROJ-12-404

In this lecture, we talked about how the person, place, and time characteristics are used in descriptive epidemiology.