1 Welcome

Infectious diseases are illnesses caused by the presence of pathogens, such as bacteria, viruses, fungi or parasites. These pathogens can infect a person, animal, or plant and can cause a wide range of symptoms, from mild to severe. Examples of infectious diseases include cholera, influenza, tuberculosis, and AIDS. Some infectious diseases can be treated with antibiotics or antiviral medications, while others have no specific treatment. Preventative measures such as vaccinations and good hygiene can help reduce the spread of infectious diseases.

This course is all about studying infectious diseases—their surveillance, their spread, their countermeasures, and their pathologies—through the lens of math and computation. There are many courses out there that focus on modeling of infectious diseases, but they often fall into one of two buckets that this course will not. The first bucket contains *plug-and-play* modeling courses which focus on using existing software to conduct simulations. The second bucket contains *theoretical* courses which focus heavily on the mathematics and analysis of infectious disease modeling equations for their own sake. Both types of courses are valuable, but this course tries to find a very practical middle ground. Instead of giving you a rental car and Driver's Ed (plug-and-play) or a dive into the thermodynamics of internal combustion engines (theoretical), I want you to emerge from this course an engineer and mechanic who understands how to put pieces together and build the vehicle you need to do the job in front of you. Buckle up.

This class is all about creating **mathematical and computational models** that help us understand when an outbreak grows, when it declines, how effective countermeasures may be, and more. We'll sometimes take detours to talk about the infectious diseases themselves, or other detours into the math or code, but we'll always come back to the underlying focus of any infectious disease modeling efforts: trying to better understand and predict infectious diseases to reduce the global burden of disease.

1.1 Pathogens vs Diseases

One thing to get used to early in learning about infectious diseases is that non-specialists often play fast and loose with the difference between a disease and the pathogen that causes it. For instance, COVID-19 is caused by the virus SARS-CoV-2. Sleeping sickness is caused by *Trypanosoma brucei*. AIDS is caused by HIV. I want you to begin to gently separate in your mind the virus, bacterium, or parasite and the disease it causes.

It's also the case that we may bundle the diseases caused by multiple pathogens into a single named disease, exemplified by malaria and influenza. Malaria is caused by parasites in the genus *Plasmodium*. This case is interesting because *P. falciparum* infection leads to markedly different disease than does *P. vivax*,[1] and indeed, both of these protozoan parasite species have different lifecycles inside the human host. Nevertheless,

[1] When we refer to genus and species, we often abbreviate the genus with a capitalized letter. For instance, you've probably heard of *Escherichia coli* as *E. coli*. Notice also that we italicize genus and species, and capitalize genus. This can sometimes be a nice hint as to whether we're looking at a pathogen or the disease it causes.

both diseases are generally called malaria. The flu is caused by the influenza virus, but there are also four types of influenza virus, labeled A, B, C and D. Both influenza A and influenza B viruses can cause severe disease in humans, while influenza C infectious are mild and non-epidemic, and influenza D doesn't infect humans to cause illness at all. In other words, the same named disease may be caused by multiple types or species of a larger genus or set of pathogens.

1.2 Koch's postulates

Koch's postulates are a set of criteria that were developed by Robert Koch in the late 19th century to establish a causal relationship between a microbe and a disease. The postulates are:

1. The microbe must be present in all cases of the disease and absent from healthy individuals.
2. The microbe must be isolated and grown in pure culture.
3. The disease must be reproduced when the pure culture is introduced into a healthy host.
4. The microbe must be re-isolated from the experimentally infected host and shown to be the same as the original culture.

Koch's postulates are still widely used today as a basic framework for identifying the causative agent of an infectious disease.[2] What Koch's postulate's don't help us with is understanding the mechanisms responsible for becoming infected in the first place. In other words, they can help us understand *what* causes an infectious disease, but not necessarily *how* it spreads.

1.3 John Snow

Cholera is caused by the bacterium *Vibrio cholerae*, which was first isolated in 1854. However, the year 1854 is more famous in the cholera story because that is the year that John Snow determined cholera's **mode of transmission**. The story is important, and Snow is sometimes considered to be one of the founders of modern epidemiology!

What did Snow do, exactly? Picture London in 1854 without sewers in some parts of the city. Without sewers, all the city's filth went into the river Thames, and from the river came drinking water. In late August of 1854, a major outbreak of cholera occurred in the Soho neighborhood and within weeks over 500 had died.

Snow made two important contributions. First, he looked at *where* the deaths were occurring, and found them to be clustered near a water pump on Broad Street. Second, hypothesizing that the water from the pump was causing people to fall ill, he removed the handle of the pump! Later, Snow created a famous map, which you can see in Fig. 1, showing the clustering of deaths around the famous Broad Street pump.

[2] When might these postulates not be fulfilled for things we think of as infectious diseases? Can you think of any reasons a disease or microbe might not check these boxes?

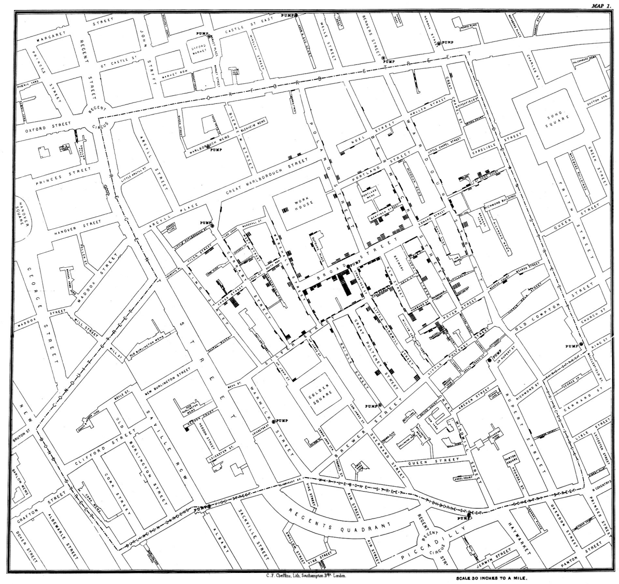

Figure 1: Snow's famous dot map of cholera deaths in the 1854 Broad Street story.

I want you to take two key lessons from John Snow's story. First, Snow illustrates the idea of a **countermeasure**, which is an intervention that decreases the spread of disease. In Snow's case, this countermeasure was based on clever reasoning and a strong hypothesis which turned out to be correct: cholera was spreading through the water, and through the Broad Street pump in particular. Today, we have lots of other countermeasures including vaccines, treatment, and more[3].

Second, however, is the more often overlooked comment from Snow[4] which essentially states that the intervention (removing the pump handle) may have simply coincided with an outbreak already in decline! In other words, cases and deaths were already naturally declining by the time Snow intervened. Why would an outbreak kill 500 in a few weeks, but then go into decline, seemingly naturally? By the end of this class, we'll have some good ideas, drawn from our understanding of mathematical models of infectious diseases.

1.4 Reservoirs and modes of transmission

A **reservoir** of disease is a place or organism where a specific disease-causing agent (such as a virus or bacteria) can survive and multiply. This can include animals (*Borrelia burgdorferi* causes Lyme disease and

[3] Can you name a few more countermeasures that you have encountered in your life?

[4] Snow: "It will be observed that the deaths either very much diminished, or ceased altogether, at every point where it becomes decidedly nearer to send to another pump than to the one in Broad street. It may also be noticed that the deaths are most numerous near to the pump where the water could be more readily obtained."

is carried by ticks and deer), plants, humans[5] or even non-living things like soil (e.g. *Clostridium tetani*; tetanus) or water (*Vibrio cholerae*; cholera). The disease-causing agent can then be transmitted from the reservoir to humans or other animals, causing illness.[6]

If a reservoir is where a disease's causal agent can survive and multiply, its **mode of transmission** is how that agent is transmitted. Epidemiologists often classify modes as direct transmission or indirect transmission, though here we should be careful about our intuitions about these words.

- **Direct transmission** means moving directly from the reservoir to a susceptible host, which includes skin-to-skin contact (gonorrhea), kissing (mononucleosis), and sexual intercourse (HIV), as well as contact with soil (hookworm) or vegetation. Direct transmission also includes droplet spread, i.e. via short-range aerosols produced while talking, singing, or coughing (pertussis, aka whooping cough).
- **Indirect transmission** means moving from the reservoir to the host with an intermediary involved. This could include inanimate objects (vehicles such as needles; HIV) or animate objects (vectors including fleas; *Yersinia pestis* aka plague). Sometimes animate objects are simply mechanical means of transmission, as in the case of fleas carrying *Yersinia pestis*, while other are biological means of transmission, as in the case of *Anopheles* mosquitos carrying *P. falciparum* malaria parasites which mature and grow inside the mosquito host.

Importantly, indirect transmission also includes airborne transmission. This is when droplets are so small that instead of falling out of the air (direct transmission) they are so small that they instead float in the air (indirect transmission). While you may be familiar with SARS-CoV-2 (COVID-19) as being spread through the air, another virus that can become airborne is *Measles morbillivirus* (MeV; measles).[7]

Because we'll mostly focus on diseases that (1) affect humans, and (2) have human or animal reservoirs in this course, this means that we'll be thinking about how humans spread infections *to each other* and how humans infect vectors and vice versa. This knowingly leaves out modeling in agriculture and the environment more broadly, including important diseases we've already mentioned such as cholera.

1.5 The need for computational and mathematical modeling of infectious diseases

Infectious diseases are all around us. In late 2019, SARS-CoV-2 emerged causing the COVID-19 pandemic, but in May 2022 we saw a global outbreak of mpox, and are still living with the HIV pandemic that started in the 1980s. Malaria killed over half a million people, mostly children, and caused over 250 million cases in 2024 alone. Measles is infecting religious communities in the U.S. and Canada, and will likely spread

[5] Can you think of any infectious diseases whose agents have a human reservoir?

[6] A famous example of shoe-leather epidemiology and reservoirs comes from the 1976 meeting of the American Legion. Before clicking the link to read more, can you guess which disease is named after this event? See if you can find notes on prevention, and think about how those relate to the disease's reservoir.

[7] If this feels a bit confusing—big droplets are "direct" while small droplets are "indirect"—that's ok! The important thing is to understand the mode of transmission, and not get too distracted by whether it is strictly considered direct or indirect when things like droplet size exist rather obviously on a continuum.

through grade schools this fall in undervaccinated communities. And this fall, over the duration of this course, many in this classroom will become infected by a rhinovirus, a coronavirus, a strain of influenza, or another respiratory pathogen. The list goes on.

The ubiquity and impact of infectious diseases provides the motivation for the kinds of questions this course hopes to answer using mathematics and computation. Can we understand the dynamics of outbreaks or steady-state spread? How many people are likely to become infected when I get sick? How should we set up systems to provide early warning signs of new pandemics or seasonal waves? And how do the answers to these questions change as we shift from a campus to a city, to a state, or to a globe? The premise of CSCI 4897/5897, and infectious disease modeling in general, is that mathematical and computational models can help us answer questions like these and can provide actionable insights that improve human health and save lives.