This is the script used for the initial recording of the indicated podcast episode. Some minor edits may have occurred during audio production.

S1E3: Models, Cholera, & COVID-19

[Street SFX]

On August 31, 1854, events began to unfold in Soho, a bustling London neighborhood, which catapulted its inhabitants into a nightmare and into the pages of history. Over the next three days, 127 people in the neighborhood would be dead, and by September 10th, over 500 residents ~~of the neighborhood~~, many of them living on Broad Street, had died from a disease whose name struck fear in the heart of everyone at that time: the disease, cholera[1].

Back then cholera was fairly new to Europe and the Americas. London had experienced a cholera outbreak recently, so everyone was aware of the devastation cholera left in its wake. Victims suffered from vomiting and diarrhea so severe that muscles cramped and organs shut down, and all the while, the afflicted were conscious and aware of what was happening. Sometimes, but very rarely, people recovered. But cholera was more often than not a death sentence that the Grim Reaper would collect on fairly quickly after you started showing symptoms. Young, old, rich, poor, powerful, or pitiful, cholera didn't care. Cholera just killed.

Cholera had actually been killing humans in India and Asia for thousands of years. Its cause remained unknown for all that time, but I'm going to give that part away here. Cholera is caused by a bacterium named *Vibrio cholerae*[2]. A particularly virulent strain of the unseen bacterium started a pandemic in the Pacific region in 1820. It briefly

subsided before a second, more aggressive wave spread across Asia, Europe, and the Americas from 1829-1833. Another outbreak in England occurred from 1848-1849 And then, for a few years, the disease which took tens of thousands of lives seemed to disappear and the pandemics ended.

The Miasma Model, which blamed bad vapors in the air as the cause of disease, was guiding scientific thought and Public Health theory when the outbreak started in Soho, and that is what makes the 1854 cholera outbreak so noteworthy[1,3]. It was this particular outbreak that brought together a physician named John Snow, who was using a new type of data, cause of death, that was being collected by a government epidemiologist and public health officer in the General Register Office of the United Kingdom named Dr. William Farr, and a neighborhood priest at St. Luke's Church named Henry Whitehead who had the essential personal connections in the neighborhood, Together they showed that the miasma model was wrong and that the spread of cholera and similar disease could be more correctly explained by a an alternative called the contagion model that said germs not bad air made people sick. The contagion model proposed that an unknown factor passing from person to person was responsible for the disease. Using the Contagion Model, Snow correctly determined that a water pump on Broad Street, a water source that served many of the ill-fated residents, was the source of the cholera infections. His work led to the Soho Council removing the pump handle, which is considered a significant moment in the history of medicine, public health, and epidemiology. It's is even remembered by the fittingly named Pump Handle Lectures given annually in Snow's honor by the John Snow Society.

[Intro Music]

I'm Phil Gibson, and welcome to BioTA. In this episode we'll focus on two important scientific concepts, systems and models. They apply to any scientific discipline, but I will be talking about them primarily in a biological context. A system is any organized biological entity composed of interacting parts that perform some biological function or process. A model is any simplified representation of that system from a scale model you can hold to a mathematical equation.

Systems-based thinking and the use of different types of models are central to the way that biologists investigate the living world, and they are definitely essential to making sense of what's going on during a disease outbreak. We'll look at scientific models, how and why scientists use them, and how they can help us learn about new pathogens (SARS-CoV-2, for example) and how to stop them.

Before we go any further, I want to acknowledge the resources that I will be drawing from. As usual, information about how to find the references I used and other resources is given at the end of the episode. But I want to particularly highlight the book *Ghost Map* by Steven Johnson. This book tells the story of the 1854 London cholera outbreak and John Snow's work to find the source of it. Everyone should read this fabulous book to get the full immersion in a thrilling telling of the tale, but for now let me give a brief summary of the highlights of the story. Let's go back to September 1, 1854.

[music fade back to Victorian SFX]

At the time of the Broad Street cholera outbreak, the prevailing medical model of what caused cholera and how it spread was bad air, or miasma[2]. Toxic vapors in the air were thought to be the cause of afflictions and clean, fresh air was the cure. Many diseases were attributed to miasma. Supporters of the miasma model also contended that another key aspect in determining whether miasma could affect someone was their inner strength and make-up. A person's freedom from vices and their overall moral quality as a human, as perceived by the standards of the time, as well as social class and race were also thought to determine whether you were likely to get different diseases from miasmas. So, in August 1854 we have an interesting mixture of biology and bias setting an unfortunate stage for the people of Soho.

Although steps had been taken to improve matters, the horrific sewage and sanitary conditions, combined with the smell of animals and industry in many parts of London, gave the miasmatists all the proof they thought they needed to explain the cause of the disease.

So where did they get this idea that bad air causes disease?

It goes back to Hippocrates' book Airs Waters and Places that described how vapors in the atmosphere cause disease. He described how swampy areas, because they are full of the scents of decay and decomposition, were full of potentially lethal air. Ancient physicians thought that if something or somewhere smells bad, it probably indicates death or something that can make you sick. That's logical. Decaying things smell. If water smells bad, you probably shouldn't drink it. If some food smells bad, it might be going bad, so don't eat it. Just think about the feeling in your stomach

when you smell something bad. You feel sick and decide to wave off that ham sandwich that's been sitting in the sun all day after one whiff. That's your primitive lizard brain protecting you. That was all the proof they needed that bad smelling air caused illness.

They thought their model made sense, even if it didn't perfectly match what they experienced in real life. For example, regardless of the smell in the air, ancient people knew not to come into contact with someone who had leprosy. They knew not to touch the same things. Still, they accepted the miasma model, even if you had to twist the interpretation of your data to make sense. Nobody questioned if their model was wrong.

For a number of reasons outlined in Ghost Map that I don't have the time to really dive into here, John Snow realized how foolish and illogical the miasma model of disease was and why its predictions did not match what was going on in terms of who got sick and who was spared. You see, Snow was also an anesthesiologist so he knew how gases, like a supposed miasma, and people interacted. They didn't affect different people in different ways, like was observed in supposed cases of miasma. Gases affect everyone pretty much the same way. From his experience, the disease outbreak simply did not match what he would expect to see if cholera spread via the air people were breathing. Everyone should get sick, not just clusters of people here and there.

Snow, like others, was beginning to apply a new model that suggested diseases were caused by contagious pathogens passed from person to person. There were many well-known instances of this. Syphilis, smallpox, typhus, and a number of other

ailments were known to be transmitted from person to person only by contact and halted by isolation—something that did not make sense under the miasma model. Yet the miasma model was held onto tightly by the medical and scientific establishment of the mid 19th century. Partly because their observations kind of fit the model, but also because they had a number of biases against the poor and other groups that led them to incorrectly interpret what was happening. They forced the data to fit the model rather than re-working the model to explain the data.

[transition]

What Snow did to build his case that cholera was caused by a waterborne pathogen and supporting the contagion model, was that he looked at where the affected individuals lived and other data. This is where William Farr comes in. As part of his job to record deaths in London, he also began including additional information about symptoms and details about their cause of death. You see, although interested in epidemiology, Farr thought the miasma model explained disease and he didn't put much stock in Snow's ideas. Snow, however, by looking at the distance from the houses of the dead to the nearest water source and talking with the clergyman, Whitehead, who provided essential details about the cholera victims, determined only individuals who drank water from that pump got cholera. Neighborhood residents who didn't drink from the pump, like workers at the local brewery who had free ale and drank water from a separate source, didn't get cholera although they breathed the same air. The water pump on Broad Street was the source of the deadly pathogen, *Vibrio cholera*. Snow eventually convinced the local officials to remove the pump handle so no one could drink from the pump.

As historian Peter Vinten-Johansen points out, here is where the story sometimes gets a little confused[4]. The story often gets twisted to say that Snow had taken a map of Soho and added dots to it to see the pattern of death that emerged and then had the pump handle removed which then stopped the disease in its tracks. This urban legend form of the story is even found on some prominent epidemiological websites. But as *The Ghost Map,* different historians, and even Snow's personal writings point out, the cholera cases had already started to decline before the pump handle was removed[1,4,5]. Yes, the water from the pump was the source of the lethal bacterial infection, but people had already stopped drinking from it and the bacterial contamination had been reduced by natural purification processes by the time the handle was removed. Snow did prepare the so-called ghost map that showed how the infections surrounded the Broad Street pump, but that was not until after the outbreak was over. Regardless, it's a great story. It's hard to beat the drama of a lone doctor collecting data and fighting the establishment to stop a deadly disease. He changes minds to his new way of interpreting the world, and helps the field of public health take a huge step forward to save lives. Hero status achieved.

Despite his efforts and solid scientific data, it took several more decades before the contagion model and germ theory that bacteria and viruses not bad air cause disease became established and replaced the miasma model. Unfortunately, though it is the sad case for Dr. Snow, that his contribution to epidemiology wasn't recognized until after his death.

[transition music]

There is another interesting aspect of the story that seems to me to be overlooked. And that has to do with the importance of models in science. We need to think very carefully about the models we use, the assumptions we make, and how we interpret the data we collect to support, modify, or reject our models.

Let's take a minute to talk about features of systems and what makes a good scientific model. To do this, I'll use a non-biological example that everyone is familiar with, weather.

[weather SFX?]
Weather is an incredibly complex system of interactions among the atmosphere, the earth, the tilt and spin of Earth on its axis, the oceans, the vegetation, geography and many other factors. Even the physical properties of water and its behavior of water as a gas, liquid, or solid are also incredibly important. Because of this complexity, atmospheric scientists and meteorologists use models that consist of diagrams, mathematical formulas, and computer programs to explain how all those factors we just mentioned work together to influence weather. They don't account for every single thing in the environment that influences weather in their different m odels, but they include what they think are the most important forces and focus their attention on them. If the model correctly accounts for the different environmental forces and how they interact, the model will allow us to predict what our weather will be like.

And this brings up something that is extremely important to remember: Regardless of the type of model, it needs to be an accurate representation of the system it describes, and if it is, it will have predictive power. The model should allow us to accurately describe and predict what happens in the natural world. The better a model

is, the more predictive power it will have about how some aspect of the system works. The better a model is, the better the model, observations, predictions and data will all fit together.

[transition music]

~~OK, we should start with two initial concepts that I will be talking about primarily in a biological context, but they apply to any scientific discipline. The concepts are those of systems and models. A system is any organized biological entity composed of interacting parts that perform some biological function. A model is a simplified representation of that system.~~

So, let's turn our attention back to models and diseases like cholera or COVID-19. It is well established that COVID-19 is caused by a virus transmitted in the air. Coughing, sneezing, talking, anything that involves exhaling droplets into the air can spread the virus if you then breathe those droplets in or get them on your hand and then touch your eyes, nose or mouth, the disease is transmitted to a new potential host. Scientists are currently trying to collect data about the how the novel coronavirus spreads population to infect individuals. Just like Farr, Snow, and Whitehead scientists are collecting all kinds of data to build a better model and figure out how to halt this particular disease that is causing a pandemic right now.

To studying how diseases spread and what effect they have on a population public health scientists and epidemiologists use what are called *demographic models* that break down a population into different categories like age, gender, race, or whatever categories the researcher thinks is important. to study for a disease. Accounting for

differences among groups can help improve accuracy of a model and provide a better picture of what is happening.

Let's say we're making a demographic model of a new disease spreading at a particular rate in a population. To keep this simple, we will assume it affects everyone the same way, and that everyone who gets infected, infects only two more people on. the day after they get infected. We can predict what the graph would look like as the disease starts with one individual (that patient zero concept you've probably heard of). We'll plot the number of infected individuals on the y-axis and time on the x-axis. At the beginning patient zero infects two new individuals. That means you'll have two new cases for a total of three infected individuals after one day. The next day, both of those new people will infect two new individuals for four new infections and seven infected individuals total. The next day, those four newly infected people infect eight more and the total number of infected individuals jumps to 15. There are twice as many new infections each day than the day before. More individuals are infected each day.

Now, consider a time in the future on day ten. There will be 1024 new cases at the end of that day, and if each of them infects two people, there will be an increase of 2,048 new cases the following day. After two weeks, starting with one individual and each infected person causing just two new infections, there will be 16,384 new cases and a total of 32,767 individuals that have caught the disease. And at three weeks. On day 21, there will be 2,097,152 new cases and 4,194,303 total cases.

You probably noticed a pattern here. Although the growth rate remained the same. you know how in our example, one person infects only two new people, as the disease

spread to more and more individuals there was a rapid increase in the *number* of new cases each day. Our graph has an overall J-shape that really takes off about day 12 showing what is called exponential growth. This pattern of exponential growth can take a local outbreak to global pandemic proportions in little time depending on how fast the disease spreads.

One type of demographic model epidemiologist use is called the Susceptible-Infectious-Recovered or SIR model. It describes a population as being composed of individuals in those three categories and it measures the rates or probabilities of individuals moving from one category to the next. This simple model can be used with those three categories or expanded to include additional categories such as reinfections or immunized individuals in the population. The SIR is a starting point for any epidemiological study of an emerging disease like COVID-19.

An important part of the SIR model is a value called R-naught, that's a capital R with a zero subscript next to it. R naught tells us the average number of people that a typical infected individual will pass the disease on to. Like in our example, using R0 in the model assumes the population has no previous exposure to the disease, so no vaccinations or immunities.

The size of R0 ~~represents and~~ summarizes a combination of factors, such as how easily the disease is transmitted, how long an infected individual is contagious, population density and things like that into one value. If R naught = 1 then an infected person will spread the disease to one other person. The disease will spread to new people but stay at a stable, low level in the population.

If R naught is less than one, then an infected individual will probably not spread it to another person, so the disease will most likely disappear from or stay at very low levels in the population. Outbreaks, epidemics and the like are probably not going to happen. But what happens if R naught is greater than one? Well, that's a big problem because now there is a possibility of an outbreak, epidemic or pandemic. Why? Because each individual can infect multiple individuals who can also affect multiple individuals so we get an exponential increase in the occurrence of the disease. How fast the spread occurs will depend on how big R naught is. In our example, R equaled 2. Imagine what happens if R naught is greater than 2. Well, if R0 is doubled to 4 in our example, you get to over 1,000,000 new case on day 10 instead of day 20. AND even if only 1% require serious medical treatment, that is 10,000 people.

Now that your curiosity is piqued, what are the R0 values of some of the viruses we're familiar with? At the low end is MERS a respiratory disease related to COVID-19, which has an R0 value of less than 1.0. Seasonal flu has an R0 of 0.9-2.1. The virus responsible for the 1918 flu pandemic had an R0 of 1.4-2.8. Measles comes in at an astounding 12-18! What about the novel coronavirus? It's R0 is currently being estimated at somewhere around 3, which is high, although not as high as some estimates of 5 early in the pandemic. Remember R0, like any statistic, has its limitations, but it gives us an idea of what is happening so we can build more accurate models and predict what will likely happen next so we can effectively plan what to do.[6,7]

If you think about the R0 and the growth model we have been using to describe its spread, you might start to realize that no population can grow forever. No virus can

spread forever. No population can grow without limit due to a lack of space or resources. That limit is called the carrying capacity and scientists include it in different demographic models to account for the fact that there is a limit on food, space, water, etc. that prevents populations from growing unchecked.

So, what in the world can possibly limit the spread of the COVID-19? Technically the capacity for its growth and spread would be the entire human population. But thinking about the carrying capacity for the virus really isn't helpful for us to plan and respond to the pandemic. Instead, what we need to think about the number of infected individuals and the capacity we have to care for the sick.

For us right now, the availability of medical resources is of critical importance. You can think of the number of hospital beds available in an ICU, the amount of PPE, the number of nurses and doctors available as the factors that determine the carrying capacity we must consider when fighting a pandemic. So even if a disease has a low infection rate, or even better a low mortality rate, what is critically important is how many individuals are infected at any given time relative to the resources we have available to care for them. If a disease spreads unchecked, you get a lot of people sick at the same time, and you may not have enough ICU beds to put them in. That overshoot of the carrying capacity leads to deaths that could have been prevented. Preventing that overshoot, keeping the number of infected individuals below the carrying capacity is the central idea behind what is called "flattening the curve".

Flattening the curve is about slowing the rate of infections so there is not exponential increase in the number of cases and keeping the number of infected individuals below our capacity to care for them. In the absence of vaccines or other preventative

measures, the only way we can slow and maybe halt the spread of the SARS-CoV-2 virus and the incidence of lethal COVID-19 cases is by efforts to reduce R0 and slow spread of the virus. ~~through increased hygiene, social distancing, and mask use so that there is not an exponential increase in the number of infected individuals.~~ Social distancing, masks, and other measures can make it more difficult for the virus to get from an infected person to a new host, which lowers R0. The disease might have a high ability to infect someone if it can get to them, but if it can't get there, no new infection.

So how do we lower R0?

Hand washing physically removes virus particles from the game. That's easy pathogen containment. The ability for some diseases to spread, particularly respiratory disease, however, is also influenced by population density. Higher density helps the virus spread, So, if there are a lot of people around, it is important to keeping a safe distance between them because it is hard for the virus to travel over those distances. With at least six feet between you and other people makes it much less likely that you'll become infected. This helps bring R 0 down.

Another simple and effective strategy for control is masks. Even if the mesh of a mask if larger than a virus, it can be small enough to catch most of the droplets of aerosol that viruses hitch a ride on when we talk, sneeze, or cough. Masks keep the number of virus particles in the air lower, making it difficult for the virus to get from person to person, and again reduces R naught. That is why the combination of social distancing and mask use works. It's not politics, it's common sense and science. Granted, we are still working out the details as we continue to improve our current model of how this

novel coronavirus spreads and how to control it, but we know that social distancing and wearing masks work to reduce spread of the virus and consequently the number of infected individuals requiring serious medical care at any one time. That can prevent overwhelming the medical system and causing deaths that could have been prevented due to unavailable beds or medical resources.

There is no doubt about it. We are still in the early stages of human interaction with this novel coronavirus. We are analyzing a lot of data and doing a lot of statistical analyses , just like John Snow did so long ago. But we're not just looking at maps of homes around a pump, these days we're looking at maps of the coronavirus genome to figure out how the fight this pathogen. We're looking at how many people are being infected, how old they are, what co-morbidities they have, the symptoms and consequences of the disease, all kinds of information. We are collecting and analyzing an amount of demographic data that would make William Farr's head spin. And although it is clear that this disease may not be as fundamentally lethal as other more common ones (but remember we are still in an early stage of learning anything about that to be certain), what matters the most right now is how fast people are becoming infected, how many people are becoming infected, and will that number overshoot our carrying capacity of resources. We know that it is likely that everyone will eventually come into contact with this disease. Current trends indicate that many will become infected and have few symptoms. I hope it stays that way. But what we absolutely do not want to happen is to have the disease spread so rapidly that our medical infrastructure is overwhelmed. If we can slow the transmission of the virus, so that there are fewer infected individuals at any one time, and therefore fewer individuals requiring serious medical care at any one time, we can make it. That's what our models (based on current data) show.

{Transition music two]

Just like with the London cholera outbreak, or any other epidemic or pandemic ranging from the Black Death and smallpox to more recent incidents of Ebola and Zika virus, there is a lot of uncertainty in the world right now. But at least we have a model that is working, and as we learn more and more, we can improve and strengthen our model. What is essential, is that we keep a critical scientific eye on what is going on and modify our models accordingly. That's the great thing about science: when we get new data that don't fit with our preferred model of how things work, we don't throw out the data. Instead, what science does is fix or throw out the model.

Let's go back to Victorian England once more to illustrate my point. In 1848, James Gillkrest[8], Inspector General of Army Hospitals in the United Kingdom wrote a book titled *Cholera gleanings : a family hand book, enabling readers of all classes to judge for themselves of the great error into which governments were unfortunately led by men looked upon as infallible guides, who very strenuously maintained the cholera to be a disease during which 'the living shall fly from the sick they should cherish.'*

Quite a title, eh. It's a short read, only 86 pages, but in this little book, Gillkrest lays out the argument that many of the leaders and governments in Europe at the time, did not do enough to control the spread and misery caused by a cholera outbreak. Gillkrist does an excellent job of describing everything that is known about the disease, and one of the points he makes over and over, is that the available medical data and even his own experiences with the disease in India didn't fit the miasma model. He argues that leaders and scientists should have done more to figure out

what was happening because their explanations using the miasma model and blaming bad vapors was clearly not working. And one more interesting thing he says is that governments dropped restrictions on isolation and allowed businesses to reopen too soon, which led to even greater spread of the disease and more misery. Sound familiar? Well, as poet and philosopher George Santayana (born just 9 years after that cholera outbreak in London) once said, "Those who cannot remember the past are condemned to repeat it."

OUTRO

So let me try to bring this all together and summarize what I hope you see as the take-home message.

Models are incredibly important in science. Whether we're talking about a scale model you can hold in your hand or a set of mathematical equations you can run through a computer, good reliable models are an essential tool for learning and investigating. But models can also have an interesting grip on us. Models like the miasma model, are often defended and supported vigorously when they are in style, regardless of how inaccurate or even harmful their base structures and assumptions are. In some instances, incomplete information can result in an inaccurate model like miasma. A willingness to adapt the model as new data come in can fix that problem. A much more insidious problem with the miasma model and others at the time is that they rested on assumptions that were based on bias and prejudice. And that is something we must always be on guard against or else science can be perverted in some awful ways. Although assumptions often get a bad name, they are an essential part of the scientific process. The assumptions behind a model gives scientists a starting point for

what they are studying. The strength of a model is, in part, due to the strength of the assumptions it is built upon. But when those assumptions are built upon biases, the model will fail to predict or explain anything with accuracy and only serve to cause more misery.

That's kind of serious, so let's end this on a more positive note. Let's consider how Snow and his immediate colleagues as well as many other scientists have made contributions by finding the problems in old models and helping us build new, better ones. And remember how Snow and others demonstrated that as we do our science, no matter what that science is, we should do everything we can to ensure that we bring together the three essential components of systems-based thinking, experience-based learning, and evidence-based conclusions as we construct and use models to explain the world around us.

out Music swell

I'm Phil Gibson, and this has been BioTA.

Terri Gibson helped with editing and co-writing duties for this episode. And special shout out to Maggie Gibson who helped with some critical background research.

Opinions expressed here are those of the author alone.

Thanks for listening, have a great day, and take very good care of your genetic material.

BioTA is a production of Under the Juniper Studios.

References and Resources Used In Episode Development

1. Johnson, S. The Ghost Map: The Story of London's Most Terrifying Epidemic--and how it Changed Science, Cities, and the Modern World. (Penguin, 2006).
2. Karamanou, M., Panayiotakopoulos, G., Tsoucalas, G., Kousoulis, A. A. & Androutsos, G. From miasmas to germs: a historical approach to theories of infectious disease transmission. Infez. Med. **20**, 58–62 (2012).
3. Bynum, W. In retrospect: On the Mode of Communication of Cholera. Nature **495**, 169–170 (2013).
4. MATRIX. The John Snow Archive and Research Companion. http://johnsnow.matrix.msu.edu/work.php?id=15-78-162.
5. Mode of Communication of Cholera(John Snow, 1855). http://www.ph.ucla.edu/epi/snow/snowbook.html.
6. News, A. B. C. What is R-naught for the COVID-19 virus and why it's a key metric for re-opening plans. ABC News https://abcnews.go.com/Health/r0-covid-19-virus-key-metric-opening-plans/story?id=70868997.

7. Pollán, M. et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. The Lancet **0**, (2020).
8. Gillkrest, J. (James), Gillkrest, J. (James), Grey, H. G. G., Great Britain. Colonial Office. Library former owner & King's College London. Cholera gleanings [electronic resource] : a family hand book, enabling readers of all classes to judge for themselves of the great error into which governments were unfortunately led by men looked upon as infallible guides, who very strenuously maintained the cholera to be a disease during which 'the living shall fly from the sick they should cherish'. (Gibraltar : Printed at the Garrison Library Press, 1848).

SOUND FX CREDITS

"Heavy thunder Strike – no Rain QUADRO.wav" CC-0 by BlueDelta of Freesound.org
"Rain, Moderate, A.wav" CC-1 by inspector of freesound.org
"Rain Thunder.mp3" CC-0 by Nimios of Freesound.org
"Trumpet Brass fanfare.wav" by ohforheavensake of Freesound.org
"Victorian Street" CC-1 by Under the Juniper Studios of AmbientMixer.com