

ROLE MODELS

The flu hit Fort Dix like clockwork every January; it had almost become a rite of passage. Most of the soldiers would go home for Christmas each year, fanning out to all corners of the United States for their winter break. They would then return to the base, well-fed and well-rested, but also carrying whichever viruses might have been going around their hometowns. If the flu was anywhere in the country, it was probably coming back with them. Life in the cramped setting of the barracks, meanwhile, offered few opportunities for privacy or withdrawal. If someone—anyone—had caught the flu back home, he was more likely than not to spread it to the rest of the platoon. You could scarcely conjure a scenario more favorable to transmission of the disease.

Usually this was no cause for concern; tens of millions of Americans catch the flu in January and February every year. Few of them die from it, and young, healthy men like David Lewis, a nineteen-year-old private from West Ashley, Massachusetts, who had returned to Fort Dix that January, are rarely among the exceptions. So Lewis, even though he'd been sicker than most of the recruits and ordered to stay in the barracks, decided to join his fellow privates on a fifty-mile march through the snow-blanketed marshlands of central New Jersey.

He was in no mood to let a little fever bother him—it was 1976, the year of the nation's bicentennial, and the country needed order and discipline in the uncertain days following Watergate and Vietnam.¹

But Lewis never made it back to the barracks: thirteen miles into the march, he collapsed and was later pronounced dead. An autopsy revealed that Lewis's lungs were flush with blood: he had died of pneumonia, a common complication of flu, but not usually one to kill a healthy young adult like Lewis.

The medics at Fort Dix had already been nervous about that year's flu bug. Although some of the several hundred soldiers who had gotten ill that winter had tested positive for the A/Victoria flu strain—the name for the common and fairly benign virus that was going around the world that year²—there were others like Lewis who had suffered from an unidentified and apparently much more severe type of flu. Samples of their blood were sent to the Center for Disease Control (CDC) in Atlanta for further testing.

Two weeks later the CDC revealed the identity of the mysterious virus. It was not a new type of flu after all but instead something altogether more disturbing, a ghost from epidemics past: influenza virus type H1N1, more commonly known as the swine flu. H1N1 had been responsible for the worst pandemic in modern history: the Spanish flu of 1918–20, which afflicted a third of humanity and killed 50 million,³ including 675,000 in the United States. For reasons of both science and superstition, the disclosure sent a chill through the nation's epidemiological community. The 1918 outbreak's earliest manifestations had also come at a military base, Fort Riley in Kansas, where soldiers were busy preparing to enter World War I.⁴ Moreover, there was a belief at that time—based on somewhat flimsy scientific evidence—that a major flu epidemic manifested itself roughly once every ten years.⁵ The flu had been severe in 1938, 1947, 1957, and 1968;⁶ in 1976, the world seemed due for the next major pandemic.

A series of dire predictions soon followed. The concern was not an immediate outbreak—by the time the CDC had positively identified

the H1N1 strain, flu season had already run its course. But scientists feared that it foreshadowed something much worse the following winter. There had never been a case, a prominent doctor noted to the *New York Times*,⁷ in which a newly identified strain of the flu had failed to outcompete its rivals and become the global hegemon: wimpy A/Victoria stood no chance against its more virulent and ingenious rival. And if H1N1 were anywhere near as deadly as the 1918 version had been, the consequences might be very bad indeed. Gerald Ford's secretary of health, F. David Mathews, predicted that one million Americans would die, eclipsing the 1918 total.⁸

President Ford found himself in a predicament. The vaccine industry, somewhat like the fashion industry, needs at least six months of lead time to know what the hip vaccine is for the new season; the formula changes a little bit every year. If they suddenly had to produce a vaccine that guarded against H1N1—and particularly if they were going to produce enough of it for the entire nation—they would need to get started immediately. Meanwhile, Ford was struggling to overcome a public perception that he was slow-witted and unsure of himself—an impression that grew more entrenched every weekend with Chevy Chase's bumbling-and-stumbling caricature of him on NBC's new hit show, *Saturday Night Live*. So Ford took the resolute step of asking Congress to authorize some 200 million doses of vaccine, and ordered a mass vaccination program, the first the country had seen since Jonas Salk had developed the polio vaccine in the 1950s.

The press portrayed the mass vaccination program as a gamble.⁹ But Ford thought of it as a gamble between money and lives, and one that he was on the right side of. Overwhelming majorities in both houses of Congress approved his plans at a cost of \$180 million.¹⁰

By summer, however, there were serious doubts about the government's plans. Although summer is the natural low season for the flu in the United States,¹¹ it was winter in the Southern Hemisphere, when flu is normally at its peak. And nowhere, from

Auckland to Argentina, were there any signs of H1N1; instead, the mild and common A/Victoria was the dominant strain again. Indeed, the roughly two hundred cases at Fort Dix remained the only confirmed cases of H1N1 anywhere in the world, and Private Lewis's the only death. Criticism started to pour in from all quarters: from the assistant director of the CDC,¹² the World Health Organization,¹³ the prestigious British medical journal *The Lancet*,¹⁴ and the editorial pages of the *New York Times*, which was already characterizing the H1N1 threat a "false alarm."¹⁵ No other Western country had called for such drastic measures.

Instead of admitting that they had overestimated the threat, the Ford administration doubled down, preparing a series of frightening public service announcements that ran in regular rotation on the nation's television screens that fall.¹⁶ One mocked the naïveté of those who refused flu shots—"I'm the healthiest fifty-five-year-old you've ever seen—I play golf every weekend!" the balding everyman says, only to be shown on his deathbed moments later. Another featured a female narrator tracing the spread of the virus from one person to the next, dishing about it in breathy tones as though it were an STD—"Betty's mother gave it to the cabdriver . . . and to one of the charming stewardesses . . . and then she gave it to her friend Dottie, who had a heart condition and died."

The campy commercials were intended to send a very serious message: *Be afraid, be very afraid*. Americans took the hint. Their fear, however, manifested itself as much toward the vaccine as toward the disease itself. Throughout American history, the notion of the government poking needles into everyone's arm has always provoked more than its fair share of anxiety. But this time there was a more tangible basis for public doubt. In August of that year, under pressure from the drug companies, Congress and the White House had agreed to indemnify them from legal liability in the event of manufacturing defects. This was widely read as a vote of no-confidence; the vaccine looked as though it was being rushed out without adequate time for

testing. Polls that summer showed that only about 50 percent of Americans planned to get vaccinated, far short of the government's 80 percent goal.¹⁷

The uproar did not hit a fever pitch until October, when the vaccination program began. On October 11, a report surfaced from Pittsburgh that three senior citizens had died shortly after receiving their flu shots; so had two elderly persons in Oklahoma City; so had another in Fort Lauderdale.¹⁸ There was no evidence that any of the deaths were linked to the vaccinations—elderly people die every day, after all.¹⁹ But between the anxiety about the government's vaccination program and the media's dubious understanding of statistics,²⁰ every death of someone who'd gotten a flu shot became a cause for alarm. Even Walter Cronkite, the most trusted man in America—who had broken from his trademark austerity to admonish the media for its sensational handling of the story—could not calm the public down. Pittsburgh and many other cities shuttered their clinics.²¹

By late fall, another problem had emerged, this one far more serious. About five hundred patients, after receiving their shots, had begun to exhibit the symptoms of a rare neurological condition known as Guillain–Barré syndrome, an autoimmune disorder that can cause paralysis. This time, the statistical evidence was far more convincing: the usual incidence of Guillain–Barré in the general population is only about one case per million persons.²² In contrast, the rate in the vaccinated population had been ten times that—five hundred cases out of the roughly fifty million people who had been administered the vaccine. Although scientists weren't positive why the vaccines were causing Guillain–Barré, manufacturing defects triggered by the rush production schedule were a plausible culprit,²³ and the consensus of the medical community²⁴ was that the vaccine program should be shut down for good, which the government finally did on December 16.

In the end, the outbreak of H1N1 at Fort Dix had been completely isolated; there was never another confirmed case anywhere in the country.²⁵ Meanwhile, flu deaths from the ordinary A/Victoria strain

were slightly below average in the winter of 1976–77.²⁶ It had been much ado about nothing.

The swine flu fiasco—as it was soon dubbed—was a disaster on every level for President Ford, who lost his bid for another term to the Democrat Jimmy Carter that November.²⁷ The drug makers had been absolved of any legal responsibility, leaving more than \$2.6 billion in liability claims²⁸ against the United States government. It seemed like every local paper had run a story about the poor waitress or schoolteacher who had done her duty and gotten the vaccine, only to have contracted Guillain–Barré. Within a couple of years, the number of Americans willing to take flu shots dwindled to only about one million,²⁹ potentially putting the nation in grave danger had a severe strain hit in 1978 or 1979.³⁰

Ford's handling of H1N1 was irresponsible on a number of levels. By invoking the likelihood of a 1918-type pandemic, he had gone against the advice of medical experts, who believed at the time that the chance of such a worst-case outcome was no higher than 35 percent and perhaps as low as 2 percent.³¹

Still, it was not clear what had caused H1N1 to disappear just as suddenly as it emerged. And predictions about H1N1 would fare little better when it came back some thirty-three years later. Scientists at first missed H1N1 when it reappeared in 2009. Then they substantially overestimated the threat it might pose once they detected it.

A Sequel to the Swine Flu Fiasco?

The influenza virus is perpetuated by birds—particularly wild seafaring birds like albatrosses, seagulls, ducks, swans, and geese, which carry its genes from one continent to another but rarely become

sick from the disease. They pass it along to other species, especially pigs and domesticated fowl like chickens,³² which live in closer proximity to humans. Chickens can become ill from the flu, but they can usually cope with it well enough to survive and pass it along to their human keepers. Pigs are even better at this, because they are receptive to both human and avian viruses as well as their own, providing a vessel for different strains of the virus to mix and mutate together.³³

The perfect incubator for the swine flu, then, would be a region in which each of three conditions held:

1. It would be a place where humans and pigs lived in close proximity—that is, somewhere where pork was a staple of the diet.
2. It would be a place near the ocean where pigs and seafaring birds might intermingle.
3. And it would probably be somewhere in the developing world, where poverty produced lower levels of hygiene and sanitation, allowing animal viruses to be transmitted to humans more easily.

This mix almost perfectly describes the conditions found in Southeast Asian countries like China, Indonesia, Thailand, and Vietnam (China alone is home to about half the world's pigs³⁴). These countries are very often the source for the flu, both the annual strains and the more unusual varieties that can potentially become global pandemics.* So they have been the subject of most of the medical community's attention, especially in recent years because of the fear over another strain of the virus. H5N1, better known as bird flu or avian flu, has been simmering for some years in East Asia and could be extremely deadly if it mutated in the wrong way.

These circumstances are not exclusive to Asia, however. The Mexican state of Veracruz, for instance, provides similarly fertile conditions for the flu. Veracruz has a coastline on the Gulf of Mexico,

and Mexico is a developing country with a culinary tradition that heavily features pork.³⁵ It was in Veracruz—where very few scientists were looking for the flu³⁶—that the 2009 outbreak of H1N1 began.³⁷

By the end of April 2009, scientists were bombarded with alarming statistics about the swine flu in Veracruz and other parts of Mexico. There were reports of about 1,900 cases of H1N1 in Mexico and some 150 deaths. The ratio of these two quantities is known as the case fatality rate and it was seemingly very high—about 8 percent of the people who had acquired the flu had apparently died from it, which exceeded the rate during the Spanish flu epidemic.³⁸ Many of the dead, moreover, were relatively young and healthy adults, another characteristic of severe outbreaks. And the virus was clearly quite good at reproducing itself; cases had already been detected in Canada, Spain, Peru, the United Kingdom, Israel, New Zealand, Germany, the Netherlands, Switzerland, and Ireland, in addition to Mexico and the United States.³⁹

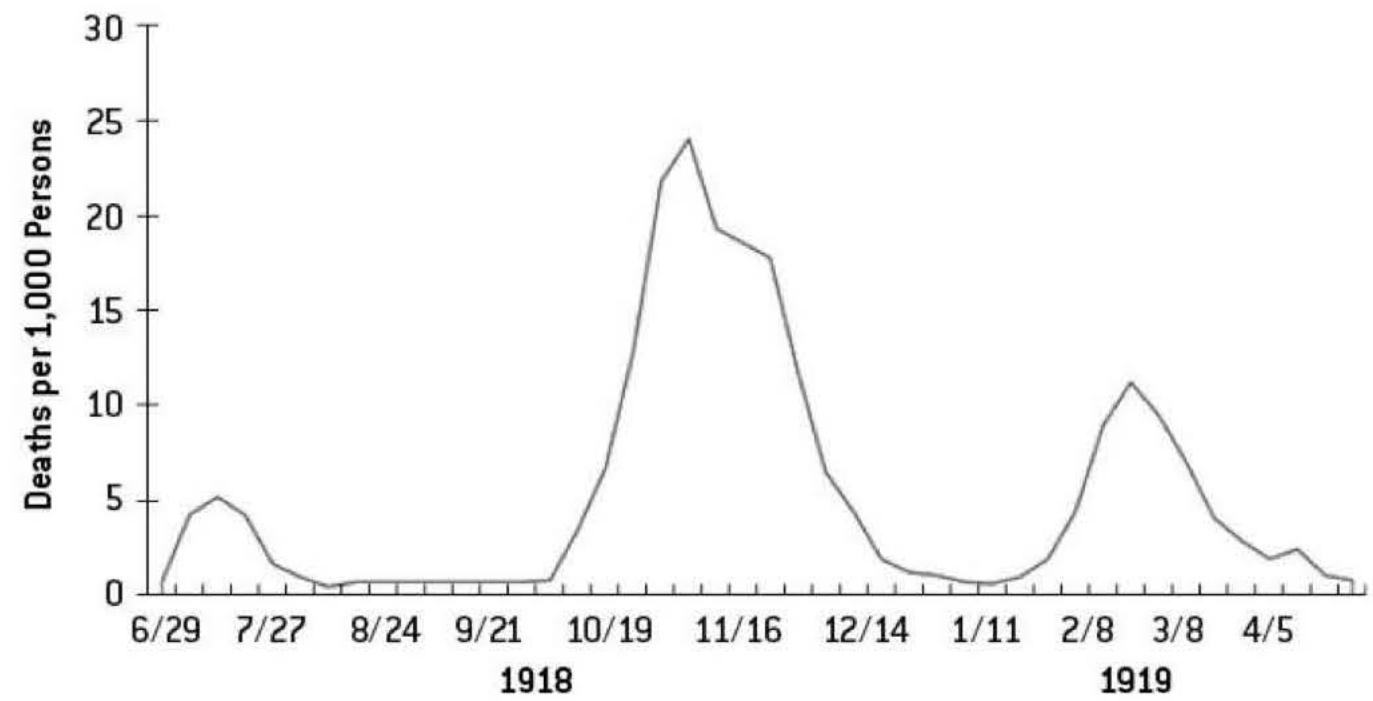
It suddenly appeared that H1N1—not H5N1—was the superbug that scientists had feared all along. Mexico City was essentially shut down; European countries warned their citizens against travel to either Mexico or the United States. Hong Kong and Singapore, notoriously jittery about flu pandemics, saw their stock markets plunge.⁴⁰

The initial worries soon subsided. Swine flu had indeed spread extremely rapidly in the United States—from twenty confirmed cases on April 26 to 2,618 some fifteen days later.⁴¹ But most cases were surprisingly mild, with just three deaths confirmed in the United States, a fatality rate comparable to the seasonal flu. Just a week after the swine flu had seemed to have boundless destructive potential, the CDC recommended that closed schools be reopened.

The disease had continued to spread across the globe, however, and by June 2009 the WHO had declared it a level 6 pandemic, its highest classification. Scientists feared the disease might follow the progression of the Spanish flu of 1918, which had initially been fairly mild, but which came back in much deadlier second and third waves

(figure 7-1). By August, the mood had again grown more pessimistic, with U.S. authorities describing a “plausible scenario” in which as much as half the population might be infected by swine flu and as many as 90,000 Americans might die.⁴²

FIGURE 7-1: DEATH RATE FROM 1918–19 H1N1 OUTBREAK



Those predictions also proved to be unwarranted, however. Eventually, the government reported that a total of about fifty-five million Americans had become infected with H1N1 in 2009—about one sixth of the U.S. population rather than one half—and 11,000 had died from it.⁴³ Rather than being an unusually severe strain of the virus, H1N1 had in fact been exceptionally mild, with a fatality rate of just 0.02 percent. Indeed, there were slightly *fewer* deaths from the flu in 2009–10 than in a typical season.⁴⁴ It hadn’t quite been the epic embarrassment of 1976, but there had been failures of prediction from start to finish.

There are no guarantees that flu predictions will do better the next time around. In fact, the flu and other infectious diseases have several

properties that make them intrinsically very challenging to predict.

The Dangers of Extrapolation

Extrapolation is a very basic method of prediction—usually, much too basic. It simply involves the assumption that the current trend will continue indefinitely, into the future. Some of the best-known failures of prediction have resulted from applying this assumption too liberally.

At the turn of the twentieth century, for instance, many city planners were concerned about the increasing use of horse-drawn carriages and their main pollutant: horse manure. Knee-deep in the issue in 1894, one writer in the *Times* of London predicted that by the 1940s, every street in London would be buried under nine feet of the stuff.⁴⁵ About ten years later, fortunately, Henry Ford began producing his prototypes of the Model T and the crisis was averted.

Extrapolation was also the culprit in several failed predictions related to population growth. Perhaps the first serious effort to predict the growth of the global population was made by an English economist, Sir William Petty, in 1682.⁴⁶ Population statistics were not widely available at the time and Petty did a lot of rather innovative work to infer, quite correctly, that the growth rate in the human population was fairly slow in the seventeenth century. Incorrectly, however, he assumed that things would always remain that way, and his predictions implied that global population might be just over 700 million people in 2012.⁴⁷ A century later, the Industrial Revolution began, and the population began to increase at a much faster rate. The actual world population, which surpassed seven billion in late 2011,⁴⁸ is about ten times higher than Petty’s prediction.

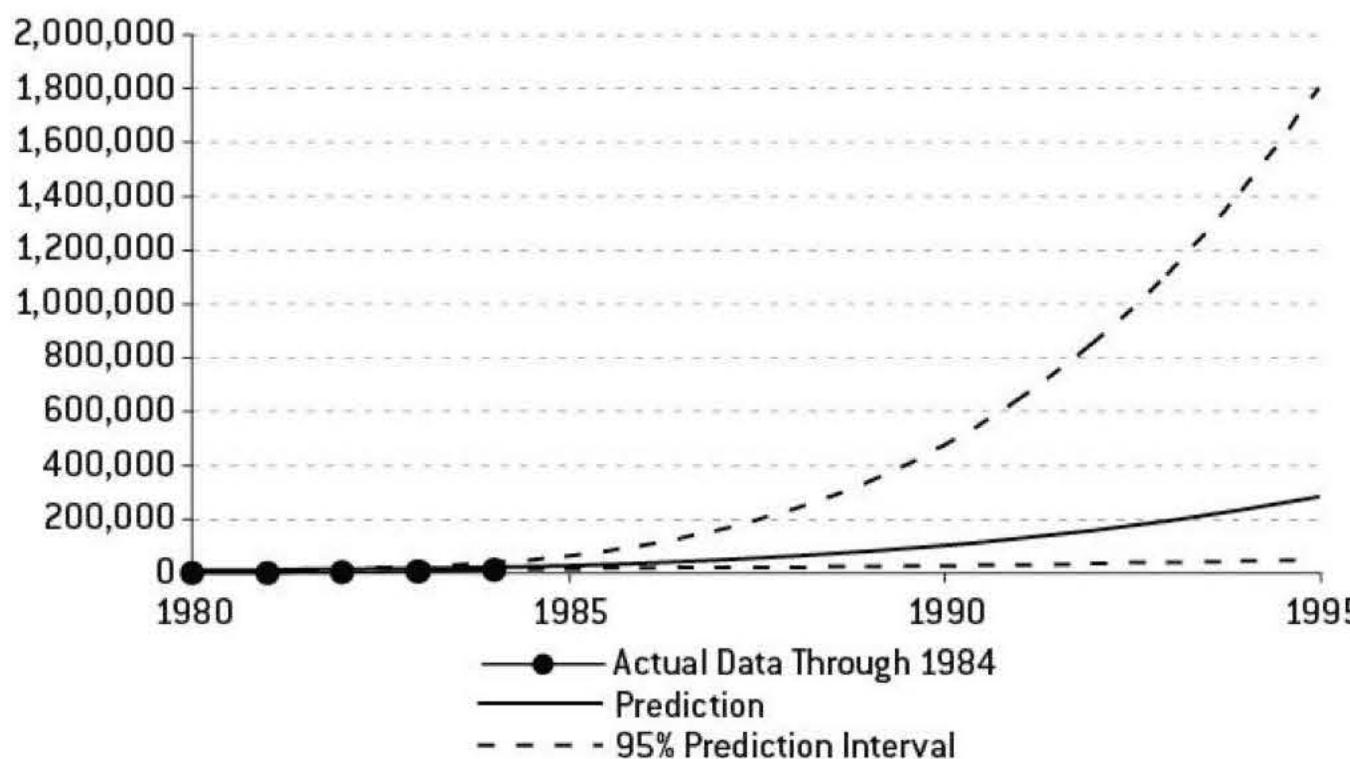
The controversial 1968 book *The Population Bomb*, by the Stanford biologist Paul R. Ehrlich and his wife, Anne Ehrlich, made the opposite mistake, quite wrongly predicting that hundreds of millions of people would die from starvation in the 1970s.⁴⁹ The reasons for this failure of prediction were myriad, including the Ehrlichs' tendency to focus on doomsday scenarios to draw attention to their cause. But one major problem was that they had assumed the record-high fertility rates in the free-love era of the 1960s would continue on indefinitely, meaning that there would be more and more hungry mouths to feed.* “When I wrote *The Population Bomb* I thought our interests in sex and children were so strong that it would be hard to change family size,” Paul Ehrlich told me in a brief interview. “We found out that if you treat women decently and give them job opportunities, the fertility rate goes down.” Other scholars who had not made such simplistic assumptions realized this at the time; population projections issued by the United Nations in the 1960s and 1970s generally did a good job of predicting what the population would look like thirty or forty years later.⁵⁰

Extrapolation tends to cause its greatest problems in fields—including population growth and disease—where the quantity that you want to study is growing exponentially. In the early 1980s, the cumulative number of AIDS cases diagnosed in the United States was increasing in this exponential fashion:⁵¹ there were 99 cases through 1980, then 434 through 1981, and eventually 11,148 through 1984. You can put these figures into a chart, as some scholars did at the time,⁵² and seek to extrapolate the pattern forward. Doing so would have yielded a prediction that the number of AIDS cases diagnosed in the United States would rise to about 270,000 by 1995. This would not have been a very good prediction; unfortunately it was too low. The actual number of AIDS cases was about 560,000 by 1995, more than twice as high.

Perhaps the bigger problem from a statistical standpoint, however, is that precise predictions aren't really possible to begin with when

you are extrapolating on an exponential scale. A properly applied version⁵³ of this method, which accounted for its margin of error, would have implied that there could be as few as 35,000 AIDS cases through 1995 or as many as 1.8 million. That's much too broad a range to provide for much in the way of predictive insight.

FIGURE 7-2: CUMULATIVE AIDS CASES DIAGNOSED IN THE UNITED STATES: ACTUAL THROUGH 1984 AND EXTRAPOLATED TO 1995



Why Flu Predictions Failed in 2009

Although the statistical methods that epidemiologists use when a flu outbreak is first detected are not quite as simple as the preceding examples, they still face the challenge of making extrapolations from a small number of potentially dubious data points.

One of the most useful quantities for predicting disease spread is a variable called the basic reproduction number. Usually designated as R_0 , it measures the number of uninfected people that can expect to

catch a disease from a single infected individual. An R_0 of 4, for instance, means that—in the absence of vaccines or other preventative measures—someone who gets a disease can be expected to pass it along to four other individuals before recovering (or dying) from it.

In theory, any disease with an R_0 greater than 1 will eventually spread to the entire population in the absence of vaccines or quarantines. But the numbers are sometimes much higher than this: R_0 was about 3 for the Spanish flu, 6 for smallpox, and 15 for measles. It is perhaps well into the triple digits for malaria, one of the deadliest diseases in the history of civilization, which still accounts for about 10 percent of all deaths in some parts of the world today.⁵⁴

FIGURE 7-3: MEDIAN ESTIMATES OF R_0 FOR VARIOUS DISEASES⁵⁵

Malaria	150
Measles	15
Smallpox	6
HIV/AIDS	3.5
SARS	3.5
H1N1 (1918)	3
Ebola (1995)	1.8
H1N1 (2009)	1.5
Seasonal flu	1.3

The problem is that reliable estimates of R_0 can usually not be formulated until well after a disease has swept through a community and there has been sufficient time to scrutinize the statistics. So

epidemiologists are forced to make extrapolations about it from a few early data points. The other key statistical measure of a disease, the fatality rate, can similarly be difficult to measure accurately in the early going. It is a catch-22; a disease cannot be predicted very accurately without this information, but reliable estimates of these quantities are usually not available until the disease has begun to run its course.

Instead, the data when an infectious disease first strikes is often misreported. For instance, the figures that I gave you for early AIDS diagnoses in the United States were those that were available only years after the fact. Even these updated statistics did a rather poor job at prediction. However, if you were relying on the data that was actually available to scientists at the time,⁵⁶ you would have done even worse. This is because AIDS, in its early years, was poorly understood and was highly stigmatized, both among patients and sometimes also among doctors.⁵⁷ Many strange syndromes with AIDS-like symptoms went undiagnosed or misdiagnosed—or the opportunistic infections that AIDS can cause were mistaken for the principal cause of death. Only years later when doctors began to reopen old case histories did they come close to developing good estimates of the prevalence of AIDS in its early years.

Inaccurate data was also responsible for some of the poor predictions about swine flu in 2009. The fatality rate for H1N1 was apparently extremely high in Mexico in 2009 but turned out to be extremely low in the United States. Although some of that has to do with differences in how effective the medical care was in each country, much of the discrepancy was a statistical illusion.

The case fatality rate is a simple ratio: the number of deaths caused by a disease divided by the number of cases attributed to it. But there are uncertainties on both sides of the equation. On the one hand, there was some tendency in Mexico to attribute deaths related to other forms of the flu, or other diseases entirely, to H1N1. Laboratory tests revealed that as few as one-quarter of the deaths supposedly linked to

H1N1 in fact showed distinct signs of it. On the other hand, there was surely underreporting, probably by orders of magnitude, in the number of cases of H1N1. Developing countries like Mexico have neither the sophisticated reporting systems of the United States nor a culture of going to the doctor at the first sign of illness;⁵⁸ that the disease spread so rapidly once the United States imported it suggests that there might have been thousands, perhaps even tens of thousands, of mild cases of the flu in Mexico that had never become known to authorities.

In fact, H1N1 may have been circulating in southern and central Mexico for months before it came to the medical community's attention (especially since they were busy looking for bird flu in Asia). Reports of an attack of respiratory illness had surfaced in the small town of La Gloria, Veracruz, in early March 2009, where the majority of the town had taken ill, but Mexican authorities initially thought it was caused by a more common strain of the virus called H3N2.⁵⁹

In contrast, swine flu was the subject of obsessive media speculation from the very moment that it entered the United States. Few cases would have gone unnoticed. With these higher standards of reporting, the case fatality rate in the United States was reasonably reliable and took some of the worst-case scenarios off the table—but not until it was too late to undo some of the alarming predictions that had been made on the public record.

Self-Fulfilling and Self-Canceling Predictions

In many cases involving predictions about human activity, the very act of prediction can alter the way that people behave. Sometimes, as in

economics, these changes in behavior can affect the outcome of the prediction itself, either nullifying it or making it more accurate. Predictions about the flu and other infectious diseases are affected by both sides of this problem.

A case where a prediction can bring itself about is called a self-fulfilling prediction or a self-fulfilling prophecy. This can happen with the release of a political poll in a race with multiple candidates, such as a presidential primary. Voters in these cases may behave tactically, wanting to back a candidate who could potentially win the state rather than waste their vote, and a well-publicized poll is often the best indication of whether a candidate is fit to do that. In the late stages of the Iowa Republican caucus race in 2012, for example, CNN released a poll that showed Rick Santorum surging to 16 percent of the vote when he had been at about 10 percent before.⁶⁰ The poll may have been an outlier—other surveys did not show Santorum gaining ground until after the CNN poll had been released.⁶¹ Nevertheless, the poll earned Santorum tons of favorable media coverage and some voters switched to him from ideologically similar candidates like Michele Bachmann and Rick Perry. Before long, the poll had fulfilled its own destiny, with Santorum eventually winning Iowa while Bachmann and Perry finished far out of the running.

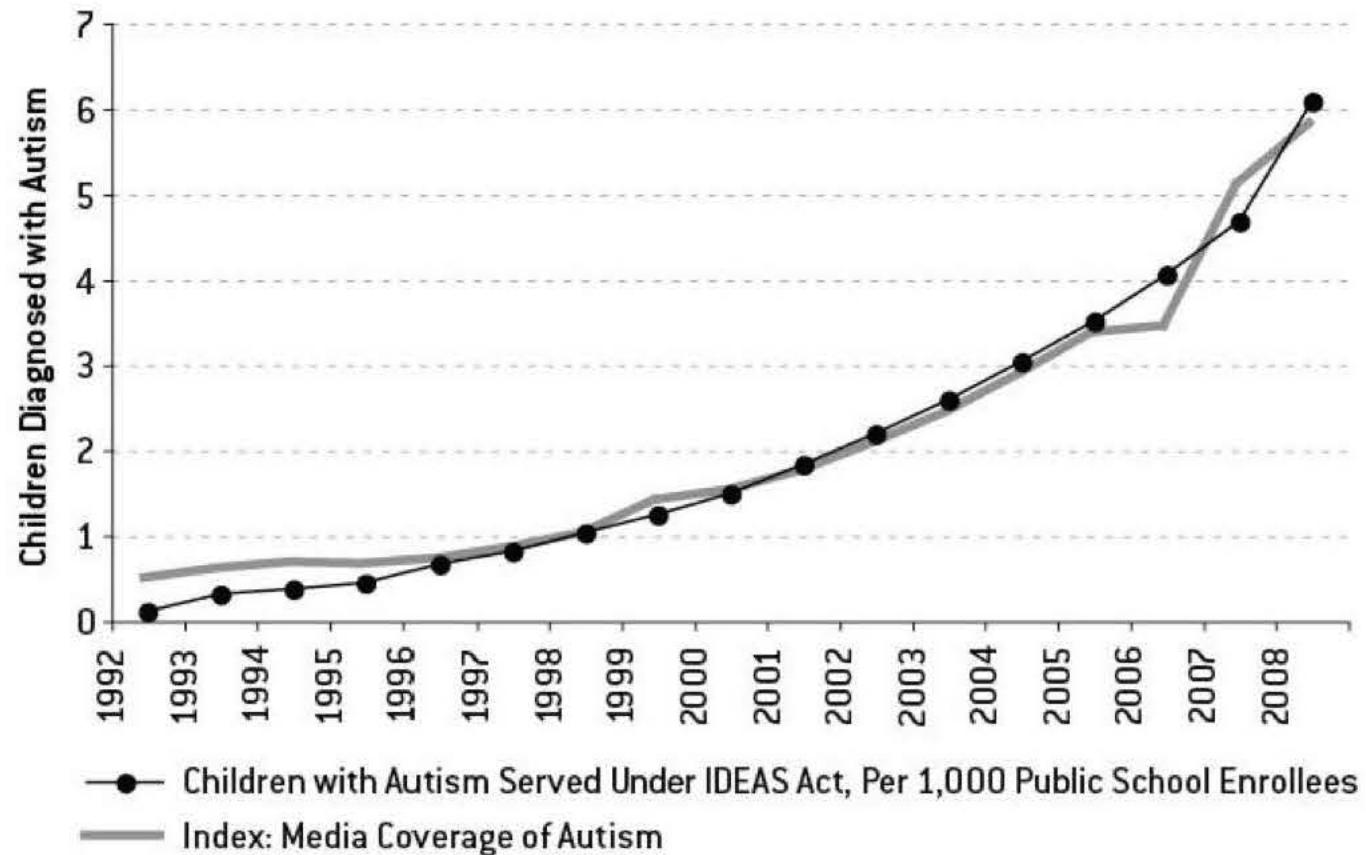
More subtle examples of this involve fields like design and entertainment, where businesses are essentially competing with one another to predict the consumer's taste—but also have some ability to influence it through clever marketing plans. In fashion, there is something of a cottage industry to predict which colors will be popular in the next season⁶²—this must be done a year or so in advance because of the planning time required to turn around a clothing line. If a group of influential designers decide that brown will be the hot color next year and start manufacturing lots of brown clothes, and they get models and celebrities to wear brown, and stores begin to display lots of brown in their windows and their catalogs, the public may well begin to comply with the trend. But they're responding more to the

marketing of brown than expressing some deep underlying preference for it. The designer may look like a savant for having “anticipated” the color, but if he had picked white or black or lavender instead, the same process might have unfolded.⁶³

Diseases and other medical conditions can also have this self-fulfilling property. When medical conditions are widely discussed in the media, people are more likely to identify their symptoms, and doctors are more likely to diagnose (or misdiagnose) them. The best-known case of this in recent years is autism. If you compare the number of children who are diagnosed as autistic⁶⁴ to the frequency with which the term *autism* has been used in American newspapers,⁶⁵ you’ll find that there is an almost perfect one-to-one correspondence (figure 7-4), with both having increased markedly in recent years. Autism, while not properly thought of as a disease, presents parallels to something like the flu.

“It’s a fascinating phenomenon that we’ve seen. In diseases that have no causal mechanism, news events precipitate increased reporting,” I was told by Dr. Alex Ozonoff of the Harvard School for Public Health. Ozonoff received his training in pure mathematics and is conversant in many data-driven fields, but now concentrates on applying rigorous statistical analysis to the flu and other infectious diseases. “What we find again and again and again is that the more a particular condition is on people’s minds and the more it’s a current topic of discussion, the closer the reporting gets to 100 percent.”

FIGURE 7-4: AUTISM—MEDIA COVERAGE AND DIAGNOSIS 1992–2008



Ozonoff thinks this phenomenon may have been responsible for some of the velocity with which swine flu seemed to have spread throughout the United States in 2009. The disease was assuredly spreading rapidly, but some of the sharp statistical increase may have come from people reporting symptoms to their doctors which they might otherwise have ignored.

If doctors are looking to make estimates of the rate at which the incidence of a disease is expanding in the population, the number of *publicly reported* cases may provide misleading estimates of it. The situation can be likened to crime reporting: if the police report an increased number of burglaries in a neighborhood, is that because they are being more vigilant and are catching crimes that they had missed before, or have made it easier to report them?^{*} Or is it because the neighborhood is becoming more dangerous? These problems are

extremely vexing for anyone looking to make predictions about the flu in its early stages.

Self-Canceling Predictions

A self-canceling prediction is just the opposite: a case where a prediction tends to undermine itself. One interesting case is the GPS navigation systems that are coming into more and more common use. There are two major north-to-south routes through Manhattan: the West Side Highway, which borders the Hudson River, and the FDR Drive, which is on Manhattan's east side. Depending on her destination, a driver may not strongly prefer either thoroughfare. However, her GPS system will tell her which one to take, depending on which has less traffic—it is predicting which route will make for the shorter commute. The problem comes when a lot of other drivers are using the same navigation systems—all of a sudden, the route will be flooded with traffic and the “faster” route will turn out to be the slower one. There is already some theoretical⁶⁶ and empirical⁶⁷ evidence that this has become a problem on certain commonly used routes in New York, Boston, and London, and that these systems can sometimes be counterproductive.

This self-defeating quality can also be a problem for the accuracy of flu predictions because their goal, in part, is to increase public awareness of the disease and therefore change the public’s behavior. The most effective flu prediction might be one that *fails* to come to fruition because it motivates people toward more healthful choices.

The Finnish scientist Hanna Kokko likens building a statistical or predictive model to drawing a map.⁶⁸ It needs to contain enough detail to be helpful and do an honest job of representing the underlying landscape—you don’t want to leave out large cities, prominent rivers and mountain ranges, or major highways. Too much detail, however, can be overwhelming to the traveler, causing him to lose his way. As we saw in chapter 5 these problems are not purely aesthetic. Needlessly complicated models may fit the noise in a problem rather than the signal, doing a poor job of replicating its underlying structure and causing predictions to be worse.

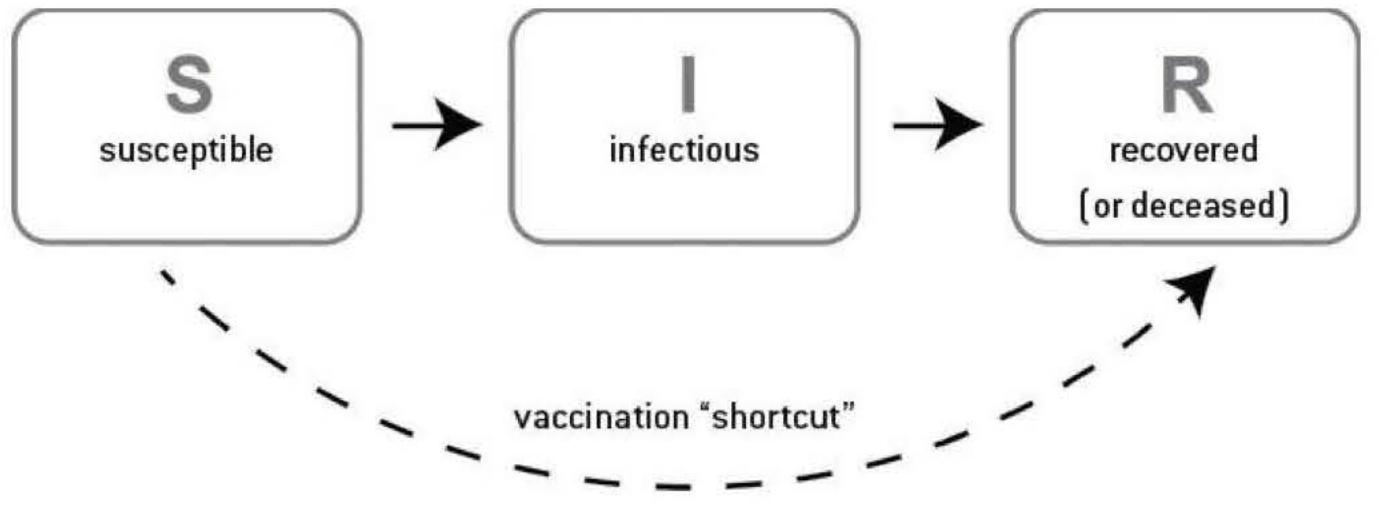
But how much detail is too much—or too little? Cartography takes a lifetime to master and combines elements of both art and science. It probably goes too far to describe model building as an art form, but it does require a lot of judgment.

Ideally, however, questions like Kokko’s can be answered empirically. Is the model working? If not, it might be time for a different level of resolution. In epidemiology, the traditional models that doctors use are quite simple—and they are not working that well.

The most basic mathematical treatment of infectious disease is called the SIR model (figure 7-5). The model, which was formulated in 1927,⁶⁹ posits that there are three “compartments” in which any given person might reside at any given time: S stands for being susceptible to a disease, I for being infected by it, and R for being recovered from it. For simple diseases like the flu, the movement from compartment to compartment is entirely in one direction: from S to I to R. In this model, a vaccination essentially serves as a shortcut,* allowing a person to progress from S to R without getting ill. The mathematics behind the model is relatively straightforward, boiling down to a handful of differential equations that can be solved in a few seconds on a laptop.

FIGURE 7-5: SCHEMATIC OF SIR MODEL

Simplicity Without Sophistication



The problem is that the model requires a lot of assumptions to work properly, some of which are not very realistic in practice. In particular, the model assumes that everybody in a given population behaves the same way—that they are equally susceptible to a disease, equally likely to be vaccinated for it, and that they intermingle with one another at random. There are no dividing lines by race, gender, age, religion, sexual orientation, or creed; everybody behaves in more or less the same way.

An HIV Paradox in San Francisco

It is easiest to see why these assumptions are flawed in the case of something like a sexually transmitted disease.

The late 1990s and early 2000s were accompanied by a marked rise in unprotected sex in San Francisco's gay community,²⁰ which had been devastated by the HIV/AIDS pandemic two decades earlier. Some researchers blamed this on increasing rates of drug use, particularly crystal methamphetamine, which is often associated with riskier sexual behavior. Others cited the increasing effectiveness of antiretroviral therapy—cocktails of medicine that can extend the lives of HIV-positive patients for years or decades: gay men no longer saw an HIV diagnosis as a death sentence. Yet other theories focused on generational patterns—the San Francisco of the 1980s, when the AIDS

epidemic was at its peak, was starting to feel like ancient history to a younger generation of gay men.⁷¹

The one thing the experts agreed on was that as unprotected sex increased, HIV infection rates were liable to do so as well.⁷²

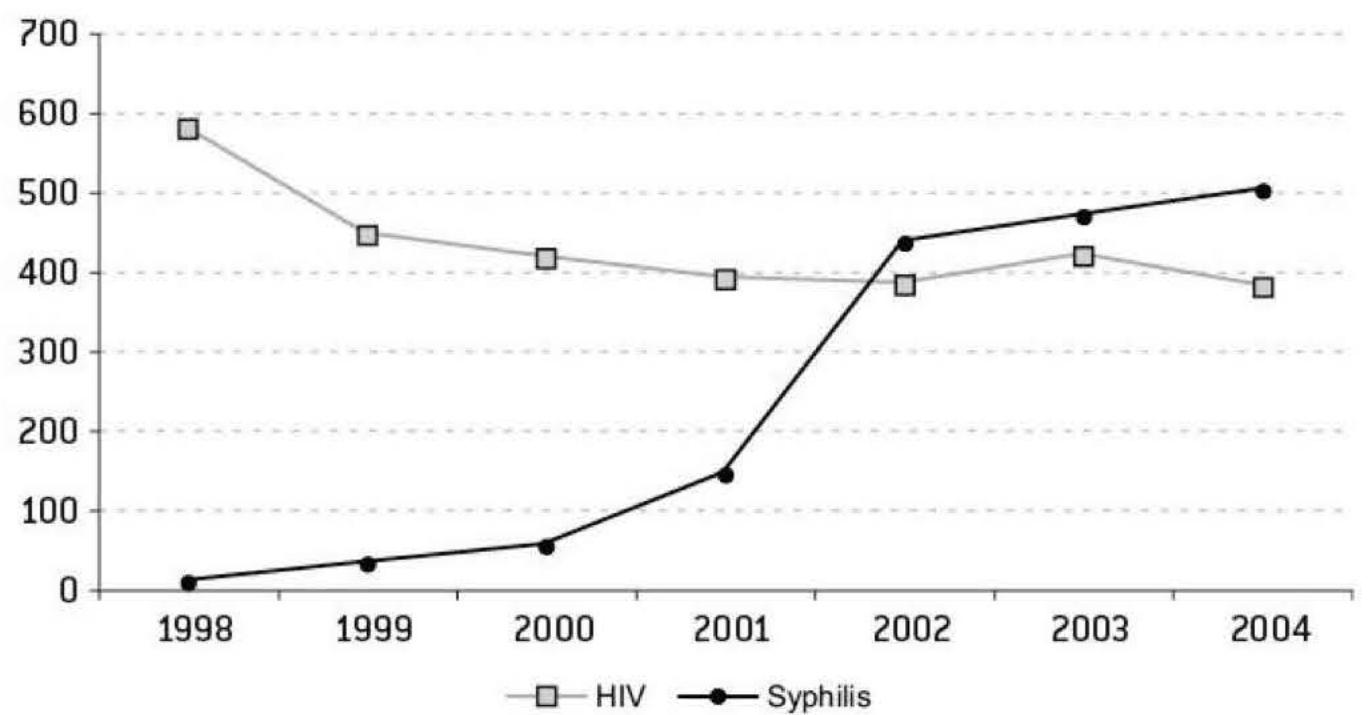
But that did not happen. Other STDs did increase: the number of new syphilis diagnoses among men who have sex with men (MSM)⁷³—which had been virtually eradicated from San Francisco in the 1990s—rose substantially, to 502 cases in 2004 from 9 in 1998.⁷⁴ Rates of gonorrhea also increased. Paradoxically, however, the number of new HIV cases did not rise. In 2004, when syphilis reached its highest level in years, the number of HIV diagnoses fell to their lowest figure since the start of the AIDS epidemic. This made very little sense to researchers; syphilis and HIV are normally strongly correlated statistically, and they also have a causal relationship, since having one disease can make you more vulnerable to acquiring the other one.⁷⁵

The solution to the paradox, it now appears, is that gay men had become increasingly effective at “serosorting”—that is, they were choosing sex partners with the same HIV status that they had. How they were able to accomplish this is a subject of some debate, but it has been documented by detailed behavioral studies in San Francisco,⁷⁶ Sydney,⁷⁷ London, and other cities with large gay populations. It may be that public health campaigns—some of which, wary of “condom fatigue,” instead focused on the notion of “negotiated safety”—were having some positive effect. It may be that the Internet, which to some extent has displaced the gay bar as the preferred place to pick up a sex partner, has different norms for disclosure: many men list their HIV status in their profiles, and it may be easier to ask tough questions (and to get honest responses) from the privacy of one’s home than in the din of the dance hall.⁷⁸

Whatever the reason, it was clear that this type of specific, localized behavior was confounding the simpler disease models—and fortunately in this case it meant that they were overpredicting HIV. Compartmental models like SIR assume that every individual has the

rate of susceptibility to a disease. That won't work as well for diseases that require more intimate types of contact, or where risk levels are asymmetric throughout different subpopulations. You don't just randomly walk into the grocery store and come home with HIV.

FIGURE 7-6: NEW DIAGNOSES OF HIV AND SYPHILIS, GAY MEN (MSM), SAN FRANCISCO, 1998–2004



How the Models Failed at Fort Dix

Even in the case of simpler diseases, however, the compartmental models have sometimes failed because of their lax assumptions. Take measles, for instance. Measles is the first disease that most budding epidemiologists learn about in their Ph.D. programs because it is the easiest one to study. "Measles is the model system for infectious disease," says Marc Lipsitch, a colleague of Ozonoff's at Harvard. "It's unambiguous. You can do a blood test, there's only one strain, and everyone who has it is symptomatic. Once you have it you don't have it again." If there's any disease that the SIR models should be able to

handle, it should be measles.

But in the 1980s and early 1990s, there were a series of unusually severe measles outbreaks in Chicago that epidemiologists were having a tough time predicting. The traditional models suggested that enough Chicagoans had been vaccinated that the population should have achieved a condition known as "herd immunity"—the biological equivalent of a firewall in which the disease has too few opportunities to spread and dies out. In some years during the 1980s, however, as many as a thousand Chicagoans—most of them young children—caught measles; the problem was so frightening that the city ordered nurses to go door-to-door to administer shots.⁷⁹

Dr. Robert Daum, a pediatrician and infectious disease specialist who works at University of Chicago hospitals, has studied these measles outbreaks in great depth. Daum is a doctor's doctor, with a dignified voice, a beard, and a detached sense of humor. He had just returned from Haiti, where he had assisted with relief efforts for the 2010 earthquake, when I visited him and two of his colleagues in Chicago.

Chicago, where I lived for thirteen years, is a city of neighborhoods. Those neighborhoods are often highly segregated, with little mixing across racial or socioeconomic lines. Daum discovered that the neighborhoods also differed in their propensity toward vaccination: inner-city residents in Chicago's mostly poor, mostly black South Side were less likely to have had their children get their MMR (measles, mumps, and rubella) shots. Those unvaccinated children were going to school together, playing together, coughing and sneezing on one another. They were violating one of the assumptions of the SIR model called random mixing, which assumes that any two members of the population are equally likely to come into contact with each other. And they were spreading measles.

This phenomenon of nonrandom mixing may also have been the culprit in the swine flu fiasco of 1976, when scientists were challenged to extrapolate the national H1N1 threat from the cases they had seen

at Fort Dix. The swine flu strain—now known as A/New Jersey/76—appeared so threatening in part because it had spread very quickly throughout the base: 230 confirmed cases were eventually diagnosed in a period of two or three weeks.⁸⁰ Thus, scientists inferred that it must have had a very high basic reproduction ratio (R_0)—perhaps as high as the 1918 pandemic, which had an R_0 of about 3.

A military platoon, however, is an usually disease-prone environment. Soldiers are in atypically close contact with one another, in cramped settings in which they may be sharing essential items like food and bedding materials, and in which there is little opportunity for privacy. Moreover, they are often undergoing strenuous physical exercise—temporarily depleting their immune systems—and the social norm of the military is to continue to report to work even when you are sick. Infectious disease has numerous opportunities to be passed along, and so transmission will usually occur very quickly.

Subsequent study⁸¹ of the outbreak at Fort Dix has revealed that the rapid spread of the disease was caused by these circumstantial factors, rather than by the disease's virulence. Fort Dix just wasn't anything like a random neighborhood or workplace somewhere in America. In fact, A/New Jersey/76 was nothing much to worry about at all—its R_0 was a rather wimpy 1.2, no higher than that of the seasonal flu. Outside a military base, or a roughly analogous setting like a college dormitory or a prison, it wasn't going to spread very far. The disease had essentially lived out its life span at Fort Dix, running out of new individuals to infect.

The fiasco over A/New Jersey/76—like the HIV/syphilis paradox in San Francisco, or the Chicago measles outbreaks of the 1980s—speaks to the limitations of models that make overly simplistic assumptions. I certainly do not mean to suggest that you should always prefer complex models to simple ones; as we have seen in other chapters in this book, complex models can also lead people astray. And because complex models often give more precise (but *not* necessarily more accurate) answers, they can trip a forecaster's sense of

overconfidence and fool him into thinking he is better at prediction than he really is.

Still, while simplicity can be a virtue for a model, a model should at least be *sophisticatedly* simple.⁸² Models like SIR, although they are useful for understanding disease, are probably too blunt to help predict its course.

SimFlu

Weather forecasts provide one of the relatively few examples of more complex models that have substantially improved prediction. It has taken decades of work, but by creating what is essentially a physical simulation of the atmosphere, meteorologists are able to do much better than purely statistical approaches to weather prediction.

An increasing number of groups are looking to apply a similar approach to disease prediction using a technique known as agent-based modeling. I visited some researchers at the University of Pittsburgh who are at the forefront of developing these techniques. The Pittsburgh team calls their model FRED, which stands for “framework for the reconstruction of epidemic dynamics.” The name is also a tip of the hat to Fred Rogers, the Pittsburgher who was the host of *Mister Rogers' Neighborhood*.

Pittsburgh, like Chicago, is a city of neighborhoods. The Pittsburgh researchers think about neighborhoods when they think about disease, and so FRED is essentially a sort of SimPittsburgh—a very detailed simulation in which every person is represented by an “agent” who has a family, a social network, a place of residence, and a set of beliefs and behaviors that are consistent with her socioeconomic status.

Dr. John Grefenstette, one of the scientists on the Pittsburgh team,

has spent most of his life in the city and still has traces of its distinct accent. He explained how FRED is organized: “They have schools and workplaces and hospitals all placed according to the right geographical distribution. They have a quite complicated setup where they assign children to schools; they don’t all go to the closest school—and some of schools are small and some of them are real large. And so you get this synthetic sort of a SimCity population.”

Dr. Grefenstette and his amiable colleague Dr. Shawn Brown showed me the results of some of FRED’s simulations, with waves of disease colorfully rippling zip code by zip code through SimPittsburgh or SimWashington or SimPhiladelphia. But FRED is also very serious business. These models take few shortcuts: literally everybody in a city, county, or state might be represented. Some agent-based models even seek to simulate the entire country or the entire world. Like weather models, they require an exponential number of calculations to be made and therefore require supercomputers to run.

These models also require a lot of data. It’s one thing to get the demographics right, which can be estimated fairly accurately through the census. But the models also need to account for human behavior, which can be much less predictable. Exactly how likely is a twenty-six-year-old Latina single mother to get vaccinated, for instance? You could draw up a survey and ask her—agent-based models rely fairly heavily on survey data. But people notoriously lie about (or misremember) their health choices: people claim they wash their hands more often than they really do, for instance,⁸³ or that they use condoms more often than they really do.⁸⁴

One fairly well-established principle, Dr. Grefenstette told me, is that people’s willingness to engage in inconvenient but healthful measures like vaccination is tied to the risk they perceive of acquiring a disease. Our SimPittsburgher will get a flu shot if she concludes that the risk from swine flu is serious, but not if she doesn’t. But how might her perception change if her neighbor gets sick, or her child does? What if there are a bunch of stories about the flu on the local news?

The self-fulfilling and self-canceling properties of disease prediction are therefore still highly pertinent to these agent-based models. Because they are dynamic and allow an agent’s behavior to change over time, they may be more capable of handling these questions.

Or consider Dr. Daum and his team at the University of Chicago, who are building an agent-based model to study the spread of a dangerous disease called MRSA, an antibiotic-resistant staph infection that can cause ordinary abrasions like cuts, scrapes, and bruises to develop into life-threatening and sometimes untreatable infections. MRSA is a complicated disease with many pathways for transmission: it can be spread through fairly casual contact like hugging, or through open wounds, or through an exchange of bodily fluids like sweat or blood. It can also sometimes linger on different types of surfaces like countertops or towels. One fairly common setting for MRSA is locker rooms, where athletes may share equipment; MRSA outbreaks have been reported among football teams ranging from high school to the NFL. Making matters more confusing still, many people carry the MRSA bacteria without ever becoming sick from it or showing any symptoms at all.

In their attempt to model MRSA, Daum and his colleagues must ask themselves questions such as these: Which types of people use a Band-Aid when they have a cut? How common is hugging in different types of cultures? How many people in a neighborhood have been to prison, where staph infections are common?

These are the sorts of questions that a traditional model can’t even hope to address, and where agent-based models can at least offer the chance of more accurate predictions. But the variables that the Pittsburgh and Chicago teams must account for are vast and wide-ranging—as will necessarily be the case when you’re trying to simulate the behavior of every individual in an entire population. Their work often takes detours into cognitive psychology, behavioral economics, ethnography, and even anthropology: agent-based models are used to study HIV infection in communities as diverse as the jungles of Papua

New Guinea⁸⁵ and the gay bars of Amsterdam.⁸⁶ They require extensive knowledge of local customs and surroundings.

Agent-based modeling is therefore an exceptionally ambitious undertaking, and the groups working in the field are often multidisciplinary All-Star teams composed of some of the best and brightest individuals in their respective professions. But for all that brainpower, their efforts are often undermined by a lack of data. “Even for H1N1, it’s been difficult to get detailed geographical data on who got sick, when, and where,” Grefenstette laments. “And it’s amazing how difficult it is to get data on past outbreaks.”

When speaking to the Pittsburgh and Chicago teams, I was sometimes reminded of the stories you read about the beautiful new shopping malls in China, which come with every imaginable frill—Roman columns, indoor roller coasters, Venetian canals—but don’t yet have any stores or people in them. Both the Chicago and Pittsburgh teams have come to a few highly useful and actionable conclusions—Dr. Grefenstette figured out, for instance, that school closures can backfire if they are too brief or occur too soon, and the U. of C. team surmised that the unusually large number of MRSA cases in inner-city Chicago was caused by the flow of people into and out of the Cook County Jail. But mostly, the models are at least a few years ahead of themselves, waiting to feed off data that does not yet exist.

The agent-based models—unlike weather forecast models that can be refined on a daily basis—are also hard to test. Major diseases come around only every so often. And even if the models are right, they might be victims of their own success because of the self-canceling property of a successful disease prediction. Suppose that the model suggests that a particular intervention—say, closing the schools in one county—might be highly effective. And the intervention works! The progress of the disease in the real world will then be slowed. But it might make the model look, in retrospect, as though it had been too pessimistic.

The Pittsburgh and Chicago teams have therefore been hesitant to

employ their models to make specific predictions. Other teams were less cautious in advance of the 2009 swine flu outbreak and some issued very poor predictions about it,⁸⁷ sometimes substantially underestimating how far the flu would spread.

For the time being the teams are mostly limited to what Dr. Daum’s colleague Chip Macal calls “modeling for insights.” That is, the agent-based models might help us to perform experiments that can teach us about infectious disease, but they are unlikely to help us predict an outbreak—for now.

How to Proceed When Prediction Is Hard

The last two major flu scares in the United States proved not to live up to the hype. In 1976, there was literally no outbreak of H1N1 beyond the cases at Fort Dix; Ford’s mass vaccination program had been a gross overreaction. In 2009, the swine flu infected quite a number of people but killed very few of them. In both instances, government predictions about the magnitude of the outbreak had missed to the high side.

But there are no guarantees the error will be in the same direction the next time the flu comes along. A human-adapted strain of avian flu, H5N1 could kill hundreds of millions of people. A flu strain that was spread as easily as the 2009 version of H1N1, but had the fatality ratio of the 1918 version, would have killed 1.4 million Americans. There are also potential threats from non-influenza viruses like SARS, and even from smallpox, which was eradicated from the world in 1977 but which could potentially be reintroduced into society as a biological weapon by terrorists, potentially killing millions. The most serious epidemics, almost by definition, can progress very rapidly: in 2009, it took H1N1

only about a week to go from a disease almost completely undetected by the medical community to one that appeared to have the potential to kill tens of millions of people.

The epidemiologists I spoke with for this chapter—in a refreshing contrast to their counterparts in some other fields—were strongly aware of the limitations of their models. “It’s stupid to predict based on three data points,” Marc Lipsitch told me, referring to the flu pandemics in 1918, 1957, and 1968. “All you can do is plan for different scenarios.”

If you can’t make a good prediction, it is very often harmful to pretend that you can. I suspect that epidemiologists, and others in the medical community, understand this because of their adherence to the Hippocratic oath. *Primum non nocere*: First, do no harm.

Much of the most thoughtful work on the use and abuse of statistical models and the proper role of prediction comes from people in the medical profession.⁸⁸ That is not to say there is nothing on the line when an economist makes a prediction, or a seismologist does. But because of medicine’s intimate connection with life and death, doctors tend to be appropriately cautious. In their field, stupid models kill people. It has a sobering effect.

There is something more to be said, however, about Chip Macal’s idea of “modeling for insights.” The philosophy of this book is that prediction is as much a means as an end. Prediction serves a very central role in hypothesis testing, for instance, and therefore in all of science.⁸⁹

As the statistician George E. P. Box wrote, “All models are wrong, but some models are useful.”⁹⁰ What he meant by that is that all models are simplifications of the universe, as they must necessarily be. As another mathematician said, “The best model of a cat is a cat.”⁹¹ Everything else is leaving out some sort of detail. How pertinent that detail might be will depend on exactly what problem we’re trying to solve and on how precise an answer we require.

Nor are statistical models the only tools we use that require us to

make approximations about the universe. Language, for instance, is a type of model, an approximation that we use to communicate with one another. All languages contain words that have no direct cognate in other languages, even though they are both trying to explain the same universe. Technical subfields have their own specialized language. To you or me, the color on the front cover of this book is yellow. To a graphic designer, that term is too approximate—instead, it’s Pantone 107.

But, Box wrote, some models are useful. It seems to me that the work the Chicago or Pittsburgh teams are doing with their agent-based models is extremely useful. Figuring out how different ethnic groups think about vaccination, how disease is transmitted throughout different neighborhoods in a city, or how people react to news reports about the flu are each important problems in their own right.

A good model can be useful even when it fails. “It should be a given that whatever forecast we make on average will be wrong,” Ozonoff told me. “So usually it’s about understanding how it’s wrong, and what to do when it’s wrong, and minimizing the cost to us when it’s wrong.”

The key is in remembering that a model is a tool to help us understand the complexities of the universe, and never a substitute for the universe itself. This is important not just when we make predictions. Some neuroscientists, like MIT’s Tomasso Poggio, think of the entire way our brains process information as being through a series of approximations.

This is why it is so crucial to develop a better understanding of ourselves, and the way we distort and interpret the signals we receive, if we want to make better predictions. The first half of this book has largely been concerned with where these approximations have been serving us well and where they’ve been failing us. The rest of the book is about how to make them better, a little bit at a time.