IF SMALLPOX STRIKES PORTLAND ...

BY CHRIS L. BARRETT, STEPHEN G. EUBANK AND JAMES P. SMITH

"EPISIMS" UNLEASHES VIRTUAL PLAGUES IN REAL CITIES TO SEE HOW SOCIAL NETWORKS SPREAD DISEASE. THAT KNOWLEDGE MIGHT HELP STOP EPIDEMICS

Suppose terrorists were to release plague in Chicago, and health officials, faced with limited resources and personnel, had to quickly choose the most effective response. Would mass administration of antibiotics be the best way to halt an outbreak? Or mass quarantines? What if a chance to nip a global influenza pandemic in the bud meant sending national stockpiles of antiviral drugs to Asia where a deadly new flu strain was said to be emerging? If the strategy succeeded, a worldwide crisis would be averted; if it failed, the donor countries would be left with less protection.

Public health officials have to make choices that could mean life or death for thousands, even millions, of people, as well as massive economic and social disruption. And history offers them only a rough guide. Methods that eradicated smallpox in African villages in the 1970s, for example, might not be the most effective tactics against smallpox released in a U.S. city in the 21st century. To identify the best responses under a variety of conditions in advance of disasters, health officials need a laboratory where "what if" scenarios can be tested as realistically as possible. That is why our group at Los Alamos National Laboratory (LANL) set out to build EpiSims, the largest individual-based epidemiology simulation model ever created.

Modeling the interactions of each individual in a population allows us to go beyond estimating the number of people likely to be infected; it lets us simulate the paths a disease would take through the population and thus where the outbreak could be intercepted most effectively. The networks that support everyday life and provide employment, transportation infrastructure, necessities and luxuries are the same ones that infectious diseases exploit to spread among human hosts. By modeling this social network in fine detail, we can understand its structure and how to alter it to disrupt the spread of disease while inflicting the least damage to the social fabric.

Virtual Epidemiology

LONG BEFORE the germ theory of disease, London physician John Snow argued that cholera, which had killed tens of thousands of people in England during the preceding 20 years, spread via the water supply. In the summer of 1854 he tested that theory during an outbreak in the Soho district. On a map, he marked the location of the homes of each of the

500 victims who had died in the preceding 10 days and noted where each victim had gotten water. He discovered that every one of them drank water from the Broad Street pump, so Snow convinced officials to remove the pump handle. His action limited the death toll to 616.

Tracing the activities and contacts of individual disease victims, as Snow did, remains an important tool for modern epidemiologists. And it is nothing new for health authorities to rely on models when developing policies to protect the public. Yet most mathematical models for understanding and predicting the course of disease outbreaks describe only the interactions of large numbers of people in aggregate. One reason is that modelers have often lacked detailed knowledge of how specific contagious diseases spread. Another is that they have not had realistic models of the social interactions in which people have contact with one another. And a third is that they have not had the computational and methodological means to build models of diseases interacting with dynamic human populations.

As a result, epidemiology models typically rely on estimates of a particular disease's "reproductive number"—the number of people likely to be infected by one contagious person or contaminated location. Often this reproductive number is a best guess based on historical situations, even though the culture, physical conditions and health status of people in those events may differ greatly from the present situation.

In real epidemics, these details matter. The rate at which susceptible people become infected depends on their individual state of health, the duration and nature of their interactions with contagious people, and specific properties of the disease pathogen itself. Truer models of outbreaks must capture the probability of disease transmission from one person to another, which means simulating not only the properties of the disease and the health of each individual but also detailed interactions between every pair of individuals in the group.

Truer models must capture the probability of disease transmission from one person to another.

Attempts to introduce such epidemiological models have, until recently, considered only very small groups of 100 to 1,000 people. Their size has been limited because they are based on actual populations, such as the residents, visitors and staff of a nursing home, so they require detailed data about individuals and their contacts over days or weeks. Computing such a large number of interactions also presents substantial technical difficulties.

Our group was able to construct this kind of individual-based epidemic model on a scale of millions of people by using high-performance supercomputing clusters and by building on an existing model called TRANSIMS developed over more than a decade at Los Alamos for urban planning [see "Unjamming Traffic with Computers," by Kenneth R. Howard; Scientific American, October 1997]. The TRANSIMS project started as a means of better understanding the potential effects of creating or rerouting roads and other transportation infrastructure. By giving us a way to simulate the movements of a large population through a realistic urban environment, TRANSIMS provided the foundation we needed to model the interactions of millions of individuals for EpiSims.

Although EpiSims can now be adapted to different cities, the original TRANSIMS model was based on Portland, Ore. The TRANSIMS virtual version of Portland incorporates detailed digital maps of the city, including representations of its rail lines, roads, signs, traffic signals and other transportation infrastructure, and produces information about traffic patterns and travel times. Publicly available data were used to generate 180,000 specific locations, a synthetic population of 1.6 million residents, and realistic daily activities for those people [*see box on opposite page*].

Integrating all this information into a computer model provides the best estimate of physical contact patterns for large human populations ever created. With EpiSims, we can release a virtual pathogen into these populations, watch it spread and test the effects of different interventions. But even without simulating a disease outbreak, the model provides intriguing insights into human social networks, with potentially important implications for epidemic response.

Overview/*Simulating Society*

- Epidemiological simulations provide virtual laboratories where health officials can test the effectiveness of different responses in advance of disease outbreaks.
- Modeling the movements of every individual in a large population produces a dynamic picture of the social network—the same network of contacts used by infectious diseases to spread among human hosts.
- Knowing the paths a disease could take through society enables officials to alter the social network through measures such as school closings and quarantines or by targeting individuals for medical treatment.

Social Networks

TO UNDERSTAND what a social network really is and how it can be used for epidemiology, imagine the daily activities and contacts of a single hypothetical adult, Ann. She has short brushes with family members during breakfast and then with other commuters or carpoolers on her way to work. Depending on her job, she might meet dozens of people

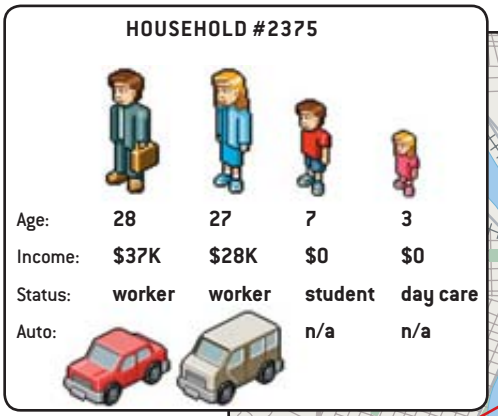
ADEPT VORMGEVING (*preceding pages*)

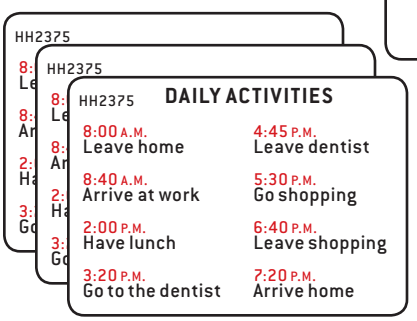
CREATING THE EPISIMS

The original EpiSims model was based on Portland, Ore., but gathering sufficiently detailed information about 1.6 million real people and their activities would have been difficult and intrusive. A synthetic population, statistically indistinguishable from the real one, could nonetheless be constructed and given realistic daily lives using publicly available data.

SYNTHETIC HOUSEHOLDS

The U.S. Census Bureau provided demographic information, such as age, household composition and income, for the entire city as well as 5 percent of its complete records for smaller study areas of a few square blocks. Through a statistical technique called iterative proportional fitting, these two data sets were combined to create households and individuals with statistically correct demographics and geographic distribution.





ACTIVITIES

Most metropolitan planning offices conduct detailed traveler activity surveys for small population samples of a few thousand. These logs track the movements of each household member over the course of one or more days, noting the time of each activity. By matching the demographics of survey respondents to the entire synthetic population, realistic daily activities can be generated for every synthetic household member.

LOCATIONS

Setting the population in motion requires assigning locations to every household's activities. Land-use data for buildings, parking lots, parks and other places were associated with 180,000 locations in the model, providing estimates of the number of people performing various types of activities there. Activities were anchored to individuals' work or school locations, and then places were chosen for additional activities, such as grocery shopping or recreation, taking into account their distance and other measures of their appeal.

at work, with each encounter having a different duration, proximity and purpose. During lunch or a shopping trip after work, Ann might have additional short contacts with strangers in public places before returning home.

We can visually represent Ann's contacts as a network with Ann in the center and a line connecting Ann to each of them [*see box on next page*]. All Ann's contacts engage in various activities and meet other people as well. We can represent these "contacts of contacts" by drawing lines from each—for example, Ann's colleague named Bob—to all his contacts. Unless they are also contacts of Ann, Bob's contacts are two "hops" away from Ann. The number of hops on the shortest path between people is sometimes called the graph distance or degree of separation between those people.

The popular idea that everyone on the earth is connected to everyone else by at most six degrees of separation means that if we continued building our social network until it included everyone on the planet, no two people would be more than six hops from one another. The idea is not strictly true, but it makes for a good story and has even led to the well-known game involving the social network of actors who have appeared in films with Kevin Bacon. In academic circles, another such social network traces mathematicians' co-authorship connections, with one's "Erdös number" defined by graph distance from the late, brilliant and prolific Paul Erdös.

Other types of networks, including the Internet, the links among scientific article citations and even the interactions among proteins within living cells, have been found to display this same tendency toward having "hubs": certain locations, people or even molecules with an unusually high number of connections to the rest of the network. The shortest path between any two nodes in the network is typically through one of these hubs, much as in a commercial airline's route system. Technically, such networks are called "scale-free" when the number of hubs with exactly k connections, $N(k)$, is proportional to a power of k [see "Scale-Free Networks," by Albert-László Barabási and Eric Bonabeau; Scientific American, May 2003].

Because a scale-free network can be severely damaged if one or more of its hubs are disabled, some researchers have extrapolated this observation to disease transmission. If infected "hub" individuals, such as the most gregarious

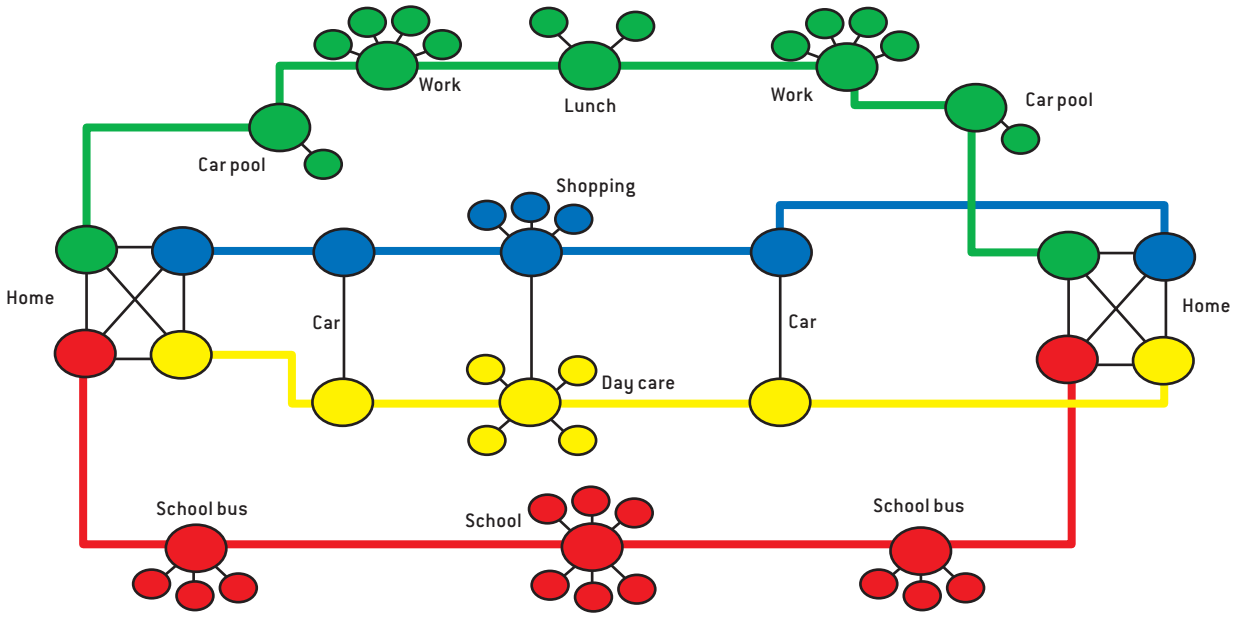
ADEPT VORMGEVING (*human figures*); LUCY READING-IKKANDA (*map*)

BUILDING SOCIAL NETWORKS

TYPICAL HOUSEHOLD'S CONTACTS

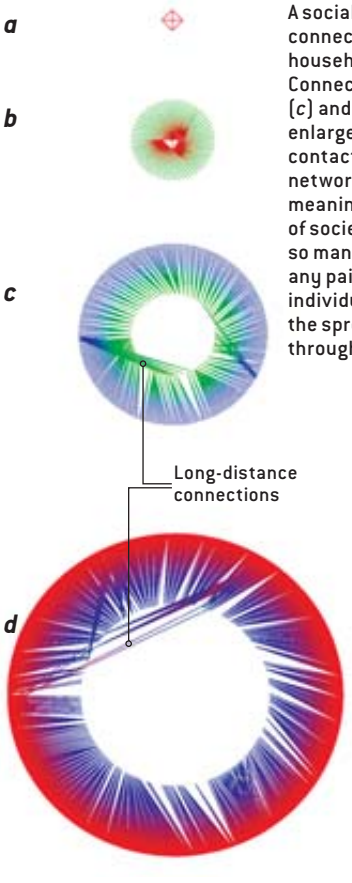
Constructing a social network for a household of two adults and two children starts by identifying their contacts with other people throughout a typical day.

This diagram shows where the household members go and what they do all day but reveals little about how their individual contacts might be interconnected or connected to others.



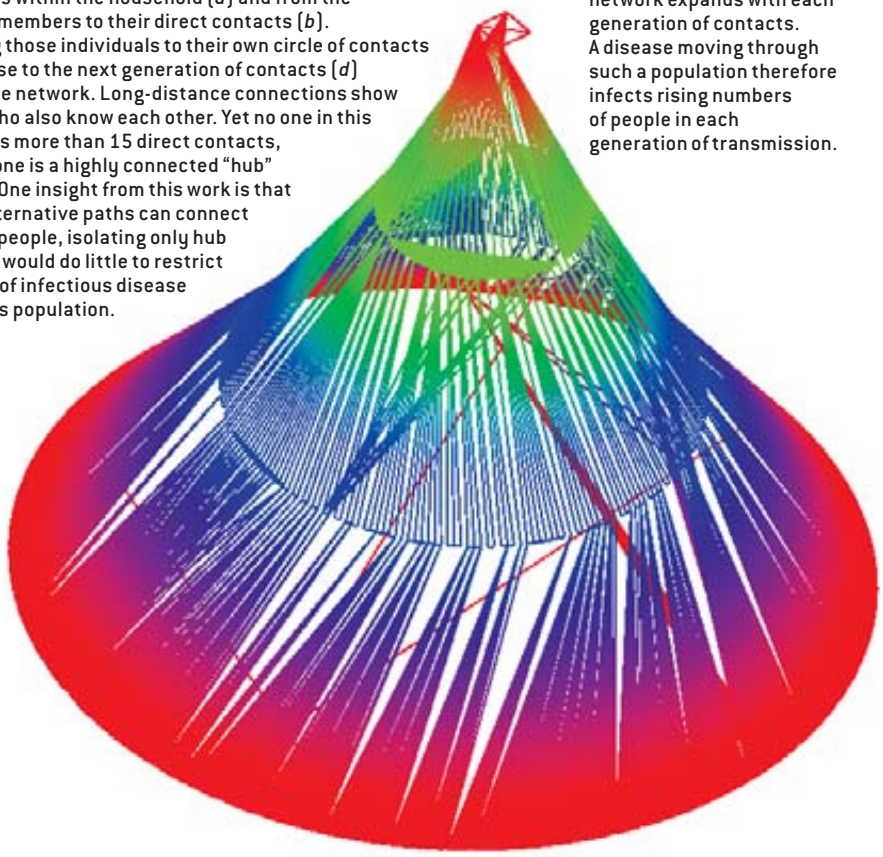
LOCAL SOCIAL NETWORK

A social network emerges by drawing lines to represent connections within the household (*a*) and from the household members to their direct contacts (*b*). Connecting those individuals to their own circle of contacts (*c*) and those to the next generation of contacts (*d*) enlarges the network. Long-distance connections show contacts who also know each other. Yet no one in this network has more than 15 direct contacts, meaning none is a highly connected "hub" of society. One insight from this work is that so many alternative paths can connect any pair of people, isolating only hub individuals would do little to restrict the spread of infectious disease through this population.



EXPANDER GRAPH

The shape of this small network expands with each generation of contacts. A disease moving through such a population therefore infects rising numbers of people in each generation of transmission.



LUCY READING-IKKANDA (*top*); STEPHEN G. EUBANK (*bottom*)

people in a population, could somehow be identified and treated or removed from the network, the reasoning goes, then an epidemic could be halted without having to isolate or treat everyone in the population. But our analyses of the social networks used by EpiSims suggest that society is not so easily disabled as physical infrastructure.

The network of physical locations in our virtual Portland, defined by people traveling between them, does indeed exhibit the typical scale-free structure, with certain locations acting as important hubs. As a result, these locations, such as schools and shopping malls, would be good spots for disease surveillance or for placing sensors to detect the presence of biological agents.

The urban social networks in the city also have human hubs with higher than average contacts, many because they work in the physical hub locations, such as teachers or sales clerks. Yet we have also found an unexpectedly high number of "short paths" in the social networks that do not go through hubs, so a policy of targeting only hub individuals would probably do little to slow the spread of a disease through the city.

In fact, another unexpected property we have found in realistic social networks is that everyone but the most devoted recluse is effectively a small hub. That is to say, when we look at the contacts of any small group, such as four students, we find that they are always connected by one hop to a much larger group. Depicting this social network structure results in what is known as an expander graph [*see box on opposite page*], which has a cone shape that widens with each hop. Its most important implication for epidemiology is that diseases can disseminate exponentially fast because the number of people exposed in each new generation of transmission is always larger than the number in the current generation.

Theoretically, this should mean that whatever health officials do to intervene in a disease outbreak, speed will be one of the most important factors determining their success. Simulating disease outbreaks with EpiSims allows us to see whether that theory holds true.

Smallpox Attack

AFTER WE BEGAN developing EpiSims in 2000, smallpox was among the first diseases we chose to model because government officials charged with bioterrorism planning and response were faced with several questions and sometimes conflicting recommendations. In the event that smallpox was released into a U.S. population, would mass vaccination be necessary to prevent an epidemic? Or would targeting only exposed individuals and their contacts for vaccination be enough? How effective is mass quarantine? How feasible are any of these options with the existing numbers of health workers, police and other responders?

Our analyses suggest that society is not so easily disabled as infrastructure.

To answer such questions, we constructed a model of smallpox that we could release into our synthetic population. Smallpox transmission was particularly difficult to model because the virus has not infected humans since its eradication in the 1970s. Most experts agree, though, that the virus normally requires significant physical contact with an infectious person or contaminated object. The disease has an average incubation period of approximately 10 days before flulike symptoms begin appearing, followed by skin rash. Victims are contagious once symptoms have appeared and possibly for a short time before they develop fever. Untreated, some 30 percent of those infected would die, but the rest would recover and be immune to reinfection.

Vaccination before exposure or within four days of infection can stop smallpox from developing. We assumed in all our simulations that health workers and people charged with tracking down the contacts of infected people had already been vaccinated and thus were immune. Unlike many epidemiological models, our realistic simulation also ensures that the chronology of contacts will be considered. If Ann contracted the disease, she could not infect her co-worker Bob a week earlier. Or, if Ann does infect Bob after she herself becomes infected and if Bob in turn infects his family member Cathy, the infection cannot pass from Ann to Cathy in less than twice the minimum incubation period between disease exposure and becoming contagious.

With our disease model established and everyone in our synthetic population assigned an immune status, we simulated the release of smallpox in several hub locations around the city, including a university campus. Initially, 1,200 people were unwittingly infected, and within hours they had moved throughout the city, going about their normal activities.

We then simulated several types of official responses, including mass vaccination of the city's population or contact tracing of exposed individuals and their contacts who could then be targeted for vaccination and quarantine. Finally, we simulated no response at all for the purpose of comparison.

In each of these circumstances, we also simulated delays of four, seven and 10 days in implementing the response after the first victims became known. In

THE AUTHORS

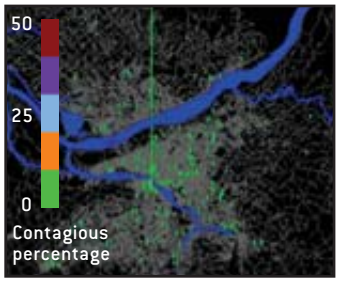
CHRIS L. BARRETT, STEPHEN G. EUBANK and JAMES P. SMITH worked for five years together at Los Alamos National Laboratory (LANL) to develop the EpiSims simulation. Barrett, who oversaw a predecessor project, TRANSIMS, is a bioinformatics specialist who now directs the Simulation Science Laboratory at the Virginia Bioinformatics Institute (VBI) in Blacksburg. Eubank, a physicist, is deputy director of the VBI simulation lab and was EpiSims team leader at Los Alamos. Smith, also a physicist, continues to work with simulations related to TRANSIMS as the project office leader for Discrete Simulation Science in the LANL Computer and Computational Sciences Division.

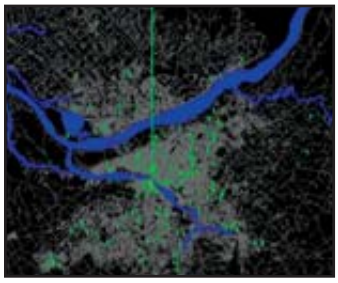
SIMULATED SMALLPOX ATTACKS

EpiSims animations depict simulated outbreaks and the effects of official interventions. In the still frames below, vertical lines indicate the number of infected people present at a location, and color shows the percentage of them who are contagious. In both scenarios shown, smallpox is released at a university in central Portland, but the attack is not detected until victims start experiencing symptoms 10 days later. The left-hand images show no public health response as a baseline. In the right-hand images, infected and exposed individuals are targeted for vaccination and quarantine. Results from a series of such simulations (*bottom*) show that people withdrawing to their homes early in an outbreak makes the biggest difference in death toll. The speed of official response, regardless of the strategy chosen, proved to be the second most important factor.

DAY 1: UNDETECTED SMALLPOX RELEASE

NO RESPONSE



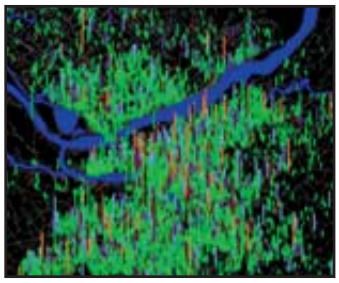


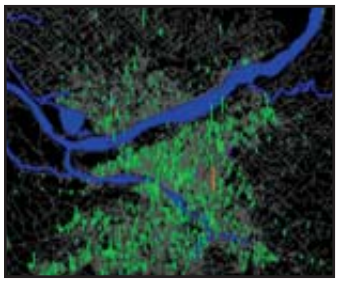
TARGETED VACCINATION AND QUARANTINE STARTING DAY 14

INFECTED: 1,281
DEAD: 0

INFECTED: 1,281
QUARANTINED: 0
VACCINATED: 0
DEAD: 0

DAY 35: SMALLPOX EPIDEMIC

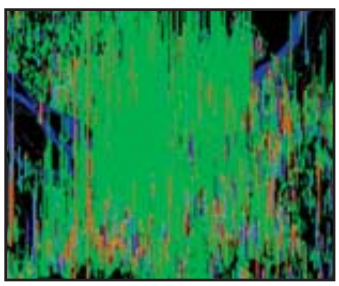


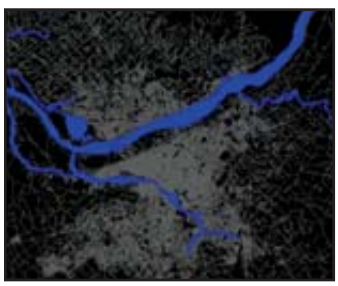


INFECTED: 23,919
DEAD: 551

INFECTED: 2,564
QUARANTINED: 29,910
VACCINATED: 30,560
DEAD: 312

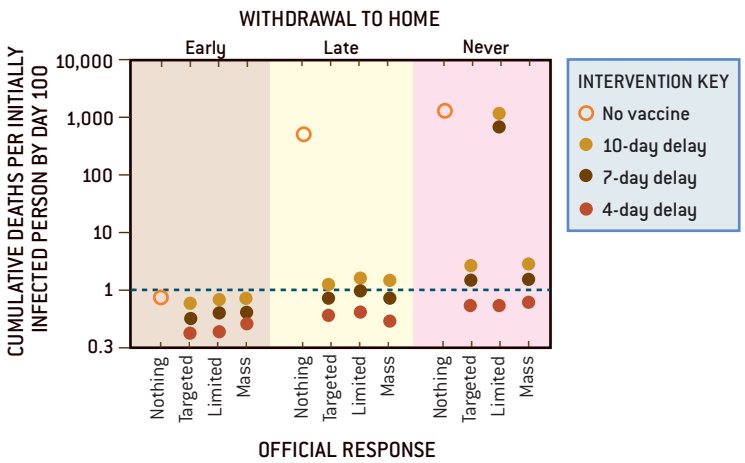
DAY 70: EPIDEMIC UNCONTAINED OR CONTAINED





INFECTED: 380,582
DEAD: 12,499

INFECTED: 2,564
QUARANTINED: 36,725
VACCINATED: 37,207
DEAD: 435



RESPONSE EFFECTIVENESS

Simulations allowed people to withdraw to their homes because they felt ill or were following officials' instructions. Withdrawal could be "early," before anyone became contagious, or "never," meaning people continued moving about unless they died. "Late" withdrawal, 24 hours after becoming contagious, was less effective than early withdrawal, which prevented an epidemic without other intervention. Official responses included doing nothing, or targeted vaccination and quarantine with unlimited personnel, or targeted vaccination limited by only half the necessary personnel being available, or mass vaccination of the entire population. The interventions began four, seven or 10 days after the first victims became symptomatic.

PHILLIP ROMERO *Los Alamos National Laboratory (simulation series)*; LUCY READING-IKKANDA *(graph)*

addition, we allowed infected individuals to isolate themselves by withdrawing to their homes.

Each simulation ran for a virtual 100 days [*see box on opposite page*], and the precise casualty figures resulting from each scenario were less important than the relative effect different responses had on the death tolls. The results upheld our theoretical prediction based on the expander-graph structure of the social network: time was by far the most important factor in limiting deaths. The speed with which people withdrew to their homes or were isolated by health officials was the strongest determinant of the outbreak's extent. The second most influential factor was the length of the delay in officials' response. The actual response strategy chosen made little difference compared with the time element.

In the case of a smallpox outbreak, these simulations indicate that mass vaccination of the population, which carries its own risks, would be unnecessary. Targeted vaccination would be just as effective so long as it was combined with rapid detection of the outbreak and rapid response. Our results also support the importance of measures such as quarantine and making sure that health officials give enforcement adequate priority during highly infectious disease outbreaks.

Of course, appropriate public health responses will always depend on the disease, the types of interventions available and the setting. For example, we have simulated the intentional release of an inhalable form of plague in the city of Chicago to evaluate the costs and effects of different responses. In those simulations we found that contact tracing, school closures and city closures each incurred economic losses of billions of dollars but did not afford many health benefits over voluntary mass use of rapidly available antibiotics at a much lower economic cost.

Most recently, as part of a research network organized by the National Institute of General Medical Sciences called the Models of Infectious Disease Agent Study (MIDAS), we have been adapting EpiSims to model a naturally occurring disease that may threaten the entire planet: pandemic influenza.

Flu and the Future

OVER THE PAST YEAR, a highly virulent strain of influenza has raged through bird populations in Asia and has infected more than 40 human beings in Japan, Thailand and Vietnam, killing more than 30 of those people. The World Health Organization has warned that it is only a matter of time before this lethal flu strain, designated H5N1, more easily infects people and spreads between them. That development could spark a global flu pandemic with a death toll reaching tens of millions [see SA Perspectives, SCIENTIFIC AMERICAN, January].

MIDAS collaborators will be studying the possibility that an H5N1 virus capable of spreading in humans might be contained or even eradicated by rapid intervention while it is still confined to a small population. To simulate the appropriate conditions in which the strain would likely emerge among humans, we are constructing a model representing a hypothetical Southeast Asian community of some 500,000 people living on farms and in neighboring small towns. Our model of the influenza virus itself will be based both on historical data about pandemic flu strains and information about the H5N1 virus, whose biology is currently a subject of intense investigation.

We know, for example, that H5N1 is sensitive to antiviral drugs that inhibit one of its important enzymes, called neuraminidase. In our simulations, we will be able to use neuraminidase inhibitors as both treatment and prophylaxis. (A vaccine against H5N1 has been developed and recently began clinical trials but because the vaccine is not yet proven or available, we will focus our simulations on seeing whether the antiviral drugs together with traditional public health measures might stop an epidemic.)

Preliminary results announced in late February are reported at www.sciam.com. In April, we will complete similar flu pandemic simulations in the EpiSims Portland model.

Our hope is that the ability to realistically model populations and disease outbreaks can help health officials make difficult decisions based on the best possible answers to "what if" questions.

The actual response chosen made little difference compared with the time element.

The creation of models such as TRANSIMS that simulate human movements through urban environments was the computational breakthrough that made EpiSims possible, and epidemiology is only one potential application for this kind of individual-based modeling. We are also in the process of creating and linking simulations of other sociotechnical systems, including environmental and atmospheric pollution, telecommunications, transportation, commodity markets, water supplies and power grids, to provide virtual laboratories for exploring solutions to a wide variety of real-world problems. SA

MORE TO EXPLORE

Scalable, Efficient Epidemiological Simulation. Stephen Eubank in *Proceedings of the 2002 ACM Symposium on Applied Computing,* pages 139–145; 2002.

Six Degrees: The Science of a Connected Age. Duncan J. Watts. W. W. Norton, 2004.

Containing Pandemic Influenza with Antiviral Agents. Ira M. Longini, Jr., et al. in *American Journal of Epidemiology,* Vol. 159, No. 7, pages 623–633; April 1, 2004.

Modelling Disease Outbreaks in Realistic Urban Social Networks. Stephen Eubank et al. in *Nature,* Vol. 429, pages 180–184; May 13, 2004.

A sample EpiSims animation and additional data from the Portland smallpox simulations can be viewed at **http://episims.lanl.gov**