

BIOSURVEILLANCE

STOPPING DISEASE BEFORE IT STOPS US

↑ Arrivals  **↑ Departures** 

Avian flu	Ebola	Hantavirus	HIV
Tuberculosis	SARS	Plague	West Nile



THIS SPRING, A NEW STRAIN OF FLU EMERGED in China, infecting poultry and causing serious disease in more than a hundred people. In the Middle East, a new coronavirus called MERS emerged last summer and by this June had killed 40 people in eight countries. The World Health Organization has warned that both viruses are particularly lethal and have the potential to cause pandemics. The world is on alert.

The good news is that flu is a well-studied enemy, and thus far, MERS is only spreading through very close contact. Scientists are expeditiously studying both pathogens, and if one does evolve to sustained human transmission, global populations will at least have a head start. However, when questioned about preparedness for new emerging diseases or bioterror attacks in general, many scientists believe that, so far, we've just been lucky.

"If SARS in 2003 had a one-to-four-day incubation period like the flu, instead of two to seven days, it would have been a completely different story," says Los Alamos biophysicist Paul Fenimore. The incubation period is the time it takes to get sick after exposure to the disease. SARS was caused by a new virus from a family that was thought to be minimally pathogenic. Identification of the virus took weeks, and the outbreak was noticed because of the pathogen's high virulence and transmissibility. If SARS had been able to spread faster than it did, many, many more lives could have been lost before scientists even knew the cause.

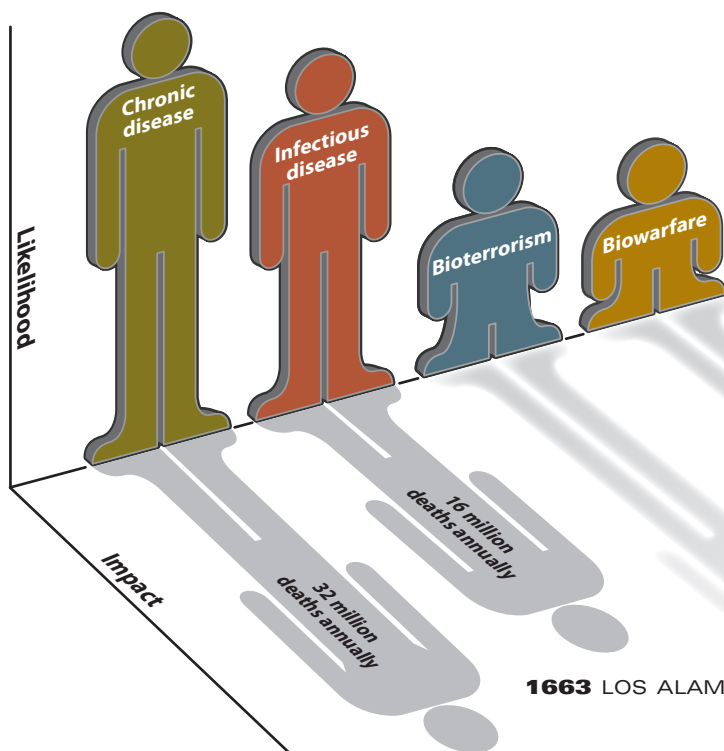
Humans have always battled contagion, but in the late 1960s, U.S. Surgeon General William H. Stewart announced it was time to "close the book on infectious disease." Sanitation, antibiotics, and vaccines had revolutionized medical care in Western society, and public health efforts began to concentrate on chronic ailments such as heart disease, cancer, and diabetes. Most remaining fear of infectious disease focused on the threat of biowarfare—prompting the ratification of the Biological Weapons Convention

in 1975, prohibiting governments from acquiring, retaining, or using biological weapons.

Yet today, infectious disease is still a major cause of death worldwide, and the rise of antibiotic resistance is bringing the threat right back to America's doorstep. Crowded cities, public transportation, and international travel and commerce make it easier for disease to spread than ever before. Furthermore, factors such as climate change and human proximity to industrial agriculture have set the stage for the emergence of new diseases (SARS), known diseases in new environments (West Nile Virus), and new versions of old diseases (drug-resistant tuberculosis). And although many governments may have agreed not to use biological weapons, individuals can use pathogens to cause widespread fear and even death. The 2001 "Amerithrax" incident, in which anthrax-containing letters were mailed to members of Congress and the media, killing five people, is a sobering example.

Interestingly, the way officials respond to an outbreak makes little distinction between anthropogenic (manmade) or naturally occurring diseases, so preparing for natural outbreaks also prepares society for bioterrorism. Part of this preparation involves biosurveillance: the process of gathering and interpreting information about disease incidence to enable a targeted response that might slow or stop its spread. Ultimately, biosurveillance attempts to predict and prevent epidemics before they start.

Los Alamos is home to many biosurveillance innovations, created by a network of more than 30 experimentalists, theorists, modelers, and engineers. Their biosurveillance "toolkit" includes disease detection, vaccine and antibiotic development, disease forecasting, response analysis, biothreat non-proliferation, and analyses of the relationship between organisms and their environments.



Preparing for natural outbreaks also prepares society for bioterrorism. Although the likelihood of bioterrorism and biowarfare may be low, the impact of such events is unknown and potentially great when considering the social, political and economic effects that would inevitably follow. In fact, there are few historical examples of bioterrorism or biowarfare, such as the U.S. Amerithrax incident in 2001 that killed five people, while annual deaths from chronic and infectious disease are comparatively predictable in both impact and likelihood. Fortunately, biological approaches to detecting pathogens, treating disease, and predicting spread are useful for all types of biothreats.



Small villages in Africa, such as this one in Uganda, are particularly hard-hit by the rise in both HIV and tuberculosis. Access to medical care is limited, and rural clinics have few resources. Reliable, portable diagnostic tests that do not require sending samples to laboratories (which are only found in larger cities) would make a substantial difference in identifying diseases quickly to reduce their spread.

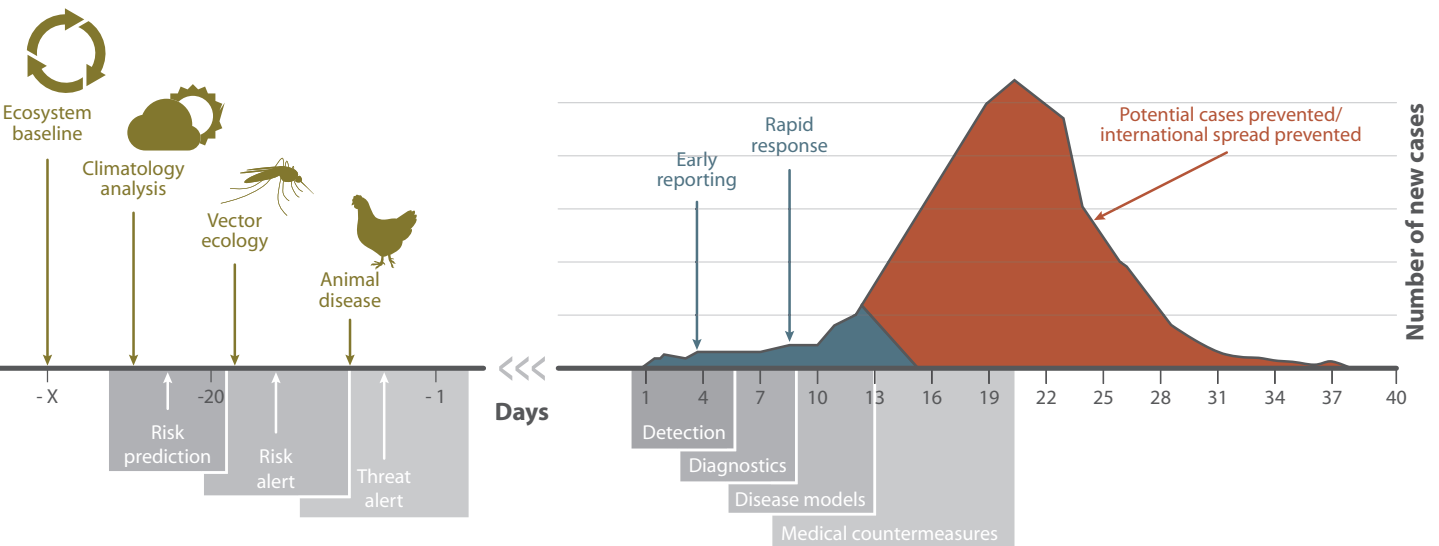
Taking it all in

Pathogens are everywhere—in plants, animals, and humans. The more that is known about what pathogens are circulating, how they are transmitted, and how they become virulent and drug-resistant, the easier it is to recognize when something goes wrong. Characterizing the circulation of pathogens worldwide means understanding human interaction with animals and the environment. To complicate this, some diseases can jump between species, some animals migrate, some diseases are vector-borne (like malaria, carried by mosquitos), and climate and weather patterns change rapidly.

“The more we know about the constant background levels of disease worldwide, the more we can understand and predict when one is going to become an epidemic,” says Jeanne Fair, an infectious disease biologist at Los Alamos. “And it’s important to learn about all disease, even when it doesn’t look like an epidemic.” For example, she notes that by understanding the percentages of viral versus bacterial respiratory illnesses in a particular area through rapid and definitive tests, local doctors could avoid unnecessary antibiotic use. Unnecessary antibiotic use leads to increased prevalence of antibiotic-resistant strains of pathogens, complicating the treatment of infected patients.

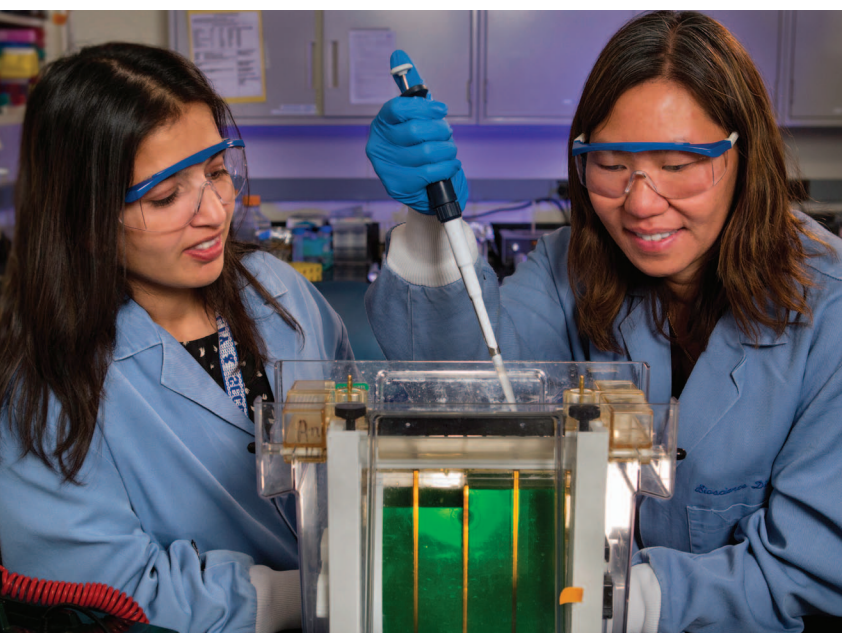
Many agencies monitor disease, including public health, defense, agricultural, and wildlife organizations; however, they do not regularly work together. In 2011, Los Alamos co-hosted a conference called *Global Biosurveillance: Enabling Science and Technology*, which was one of many international forums to share ideas on how to improve biosurveillance. Among the outcomes was a desire to integrate national and international agencies into a cohesive network to establish baselines for ecosystem risks and threats while enabling data sharing for improved surveillance and response—much in the same way that, after 9/11, authorities recognized that it would be beneficial to share information between the CIA, FBI, and local law enforcement.

A further observation from the conference was that improvements are needed to help identify disease more quickly and accurately—and without the need for extensive lab equipment so that it can be done in all parts of the world. In 2012, President Obama issued the *National Strategy for Biosurveillance*, which echoed many of these points, calling for integration, partnership, and innovation.



This timeline of events in a hypothetical disease outbreak shows that detection and diagnosis in the first week, considered early reporting, can enable a rapid response that would significantly reduce the overall spread of the disease. However, looking backwards, there are opportunities to predict coming diseases prior to the first reported case based on broad categories of information such as ecosystem and climate data, vector ecology data (e.g., mosquito populations), and diseases in animal populations.

CREDIT: ADAPTED FROM THE WORLD HEALTH ORGANIZATION'S 2007 WORLD HEALTH REPORT AND A GRAPHIC BY BILL HUFF/DOD



Harshini Mukundan (left) watches as fellow Los Alamos bioscientist Elizabeth Hong-Geller loads a vertical electrophoresis gel, used to separate proteins and small molecules. This technique is useful for developing improved methods to detect pathogens.

Invisible enemy

The first step of biosurveillance is detection: where are people sick and what diseases do they have? Diagnostic tests used by doctors at medical clinics and hospital emergency rooms are often the best source of this information. When a person is ill, the pathogen is busy replicating itself while the body's immune system is likely launching a counterattack. Most laboratory tests target these two events, and scientists at Los Alamos are working to improve both types of tests: those that target the pathogen itself in samples of blood, urine, or sputum (mucus) as well as those that recognize the body's immune response to current and previous infections by detecting antibodies (molecules that bind to invaders).

One Los Alamos team, including biochemists Basil Swanson and Harshini Mukundan, has been specializing in the diagnosis of tuberculosis (TB) for over a decade. TB is a leading cause of death in individuals with HIV/AIDS. Parts of Africa are now struggling with epidemics of both diseases.

"One of the biggest problems is that many of the current TB tests are likely to fail if the person is also infected with HIV," says Mukundan. "The tests come back negative, so people who are co-infected are being sent home without anyone realizing they have TB and without the proper care for the disease. This is partly why it continues to spread."

Commonly used methods to detect TB include a skin test (which can produce a false-positive in patients who

already have antibodies from a previous infection or from the TB vaccine), a sputum test (which is laboratory intensive, requires a highly contagious sample, and does not work for all types of TB), and a blood test (which requires costly laboratory equipment not available in all countries). Unfortunately, it is not entirely understood why some of these tests fail in HIV patients.

Making use of a recent Los Alamos invention—an optical biosensor that can detect multiple kinds of pathogens—the TB team recently developed novel strategies to detect very small concentrations of a tiny sugar called lipoarabino-mannan (LAM) that comes from the membranes of TB-causing bacteria. LAM is a virulence factor—a molecule that reveals the ability to cause disease—secreted by the bacteria, making it a useful biomarker for indicating of the presence of the TB pathogen. The team has also developed assays for two other biomarkers that together allow for a reliable diagnosis of active TB infection within minutes. Their ultimate goal is to create simple, reliable methods of detecting HIV, active TB, and other diseases in rural settings worldwide—in humans and any animal populations of interest.

Another TB project at Los Alamos focuses on the effectiveness of antibiotics. With the rising prevalence of drug-resistant TB, many in the field have suggested that perhaps the mutations that make the bacteria resistant also reduce their ability to spread. But recent work by Los Alamos biologists Bette Korber, Karina Yusim, and Shihai Feng, in collaboration with the National Institutes of Health, indicates otherwise. Their work shows that compensatory mutations can restore the fitness of the drug-resistant bacteria (that cause TB) and has confirmed the persistence and spread of drug-resistant organisms in the population.

Clues from the blueprint

A number of studies at Los Alamos are examining the complex relationship between the host organism and the pathogen, as well as the molecular blueprints (DNA and RNA) of pathogens, in order to create detection strategies. Biologist Elizabeth Hong-Geller has been examining small RNA (sRNA) molecules produced by bacteria that are involved in gene regulation during infection. Her work has focused on the bacteria *Yersinia pestis*, which causes plague. In collaboration with Lab colleagues who determine 3D biomolecular structures and create molecular models, she is trying to identify small molecules that can bind to, and potentially inhibit, key sRNAs for antibiotics and drug design.

"If a small RNA is involved in virulence and we can block its function, it would be a breakthrough for designing countermeasures against infection," says Hong-Geller.



Social Media
Diagnostic Data
Transmission Factors
Pathogen Lifecycle

Also in the genetics arena at Los Alamos, computational biologists Murray Wolinsky, Jason Gans, and Jian Song are experts at developing algorithms to find genetic signatures: unique sections of DNA or RNA that can be used to distinguish quickly between pathogens—especially closely related pathogens.

Once a signature is identified, primers made of short sections of DNA are developed using a complementary sequence so that identified regions of the pathogens' DNA will specifically bind to the primers. Biologist Norman Doggett has helped develop rapid tests called assays that screen for many types of pathogens at once by introducing sample material (serum or urine that might contain DNA or RNA from a pathogen) to multiple primers, and amplifying, or copying, the ones that find a match.

These types of assays are also great for evaluating environmental samples, such as soil or air. After the Amerithrax incident, the U.S. government began routinely monitoring the air for dangerous pathogens in major cities through a program called BioWatch. Los Alamos—which had already been involved in the analysis of the anthrax used in the letters—stepped in with expertise in analyzing BioWatch samples and optimizing the placement of detectors. There has been a lot of public scrutiny of the BioWatch program, mostly about the possibility of false positives, prompting rigorous assay validation in which Los Alamos also played a key role. The overarching problem remains: it all comes back to the sensitivity and specificity of detection methods.

For example, a detector could test positive for *Bacillus anthracis*, the bacteria that causes anthrax, when really the sample contains *Bacillus thuringensis*, a non-deadly close relative of *Bacillus anthracis*. Both bacteria live naturally in the soil and are genetically similar, but only one is a major threat to humans. Scientists have extensively studied the differences between anthrax near-neighbors and have developed discerning tests using signatures that target only the small differences in their genetic codes that account for their

pathogenicity, or ability to cause disease. However, a less well-studied organism may have unknown near-neighbors from which it would be difficult to distinguish.

The gold standard for comparing various pathogen strains is to sequence the entire genome. Over the years, genomic sequence data has been amassed in pathogen databases at Los Alamos to aid in the comparative analysis of many viruses, including influenza, HIV, and hepatitis C virus (HCV). During the 2009 swine flu episode, Los Alamos scientists were able to use the influenza database to quickly determine that the culprit was indeed a new strain. Fortunately, the diminishing cost of sequencing is enabling more organisms to be sequenced, thus generating enough information for comprehensive comparative analysis.

New techniques that enable sequencing entire communities of organisms at once (metagenomics) or sequencing only genes that are being expressed (transcriptomics)—which can change with environmental conditions—are also giving scientists much more information about pathogens. So much data, however, can sometimes be a problem. Numerous redundancies in closely related strains of organisms makes comparisons difficult. To confront this issue, bioinformaticist Patrick Chain and his team at Los Alamos have been developing a database containing only the unique sections of each organism's genetic code.

"We have been developing methods to essentially screen all known genomes for any identical sequences, track where they are in each genome, and remove them such that they will no longer confound searches for similarities between sequences," says Chain.

Overall, no matter what the approach on a molecular level, detection strategies for biosurveillance have the same goal in mind: simple, rapid, "field-able" methods. Many research projects at Los Alamos have taken on this challenge over the years. In fact, spin-off companies were created around some of these technologies such as a dipstick test (much like a pregnancy test) for the flu and a small, portable flow cytometer that uses sound waves for cell sorting.

#GotFlu

Whether or not patients receive a definitive diagnostic test at their doctor visit, notes about their symptoms are always recorded. “Syndromic surveillance” describes the idea of screening hospital and clinic records in search of trends or anomalies in patient complaints—prior to diagnostic tests—that might foretell an epidemic or biothreat event. For instance, multiple patients in the month of October complaining of upper respiratory disease with a cough, high fever, and muscle aches may suggest to a doctor that it is the beginning of flu season, even though the doctor may not perform a definitive test on each patient.

This approach, however, still requires someone to be sick enough to go to a doctor. Is it possible to detect disease prior to this point? What else do people do when they are feeling under the weather? Purchase over-the-counter drugs, Google their symptoms, or complain to their friends on Twitter? All of these actions produce potentially useful biosurveillance data.

Los Alamos biomedical scientist Alina Deshpande leads a research project to analyze all the possible data streams that could be useful for biosurveillance. Her team is studying the relevance of various data streams and developing a systematic approach to determine which data types are useful for which purposes.

To achieve this, the team evaluated many currently available sources of data (emergency room and other clinic records, social media, Internet search queries, laboratory records, etc.) for their utility, using criteria such as timeliness, granularity, and credibility. This was done using a commercially available Multi Criteria Decision Analysis (MCDA) software tool that scores data streams based on weighted metrics and assigned values specific to data stream categories, such as early detection or consequence management.

Deshpande’s team also evaluated historical outbreaks, such as the 2009 swine flu pandemic and the 2010 cholera outbreak in Haiti, to find surveillance windows, or points in time at which early detection or early warning could have made a difference. The team then researched what data streams were available at that time to determine which ones would have been useful. A cross-method analysis was performed between the surveillance window evaluation and the MCDA evaluation to identify data stream categories that

One of the biggest challenges in biosurveillance is determining how to capture and integrate all the useful data streams into actionable information.

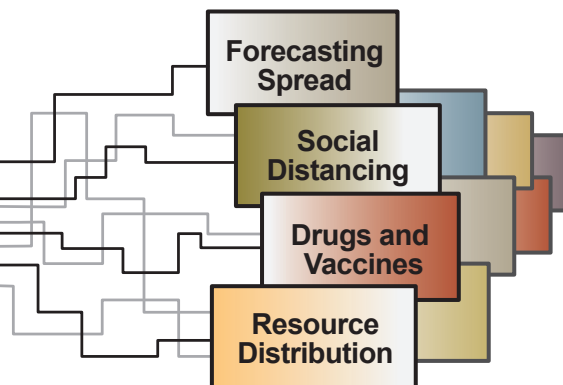


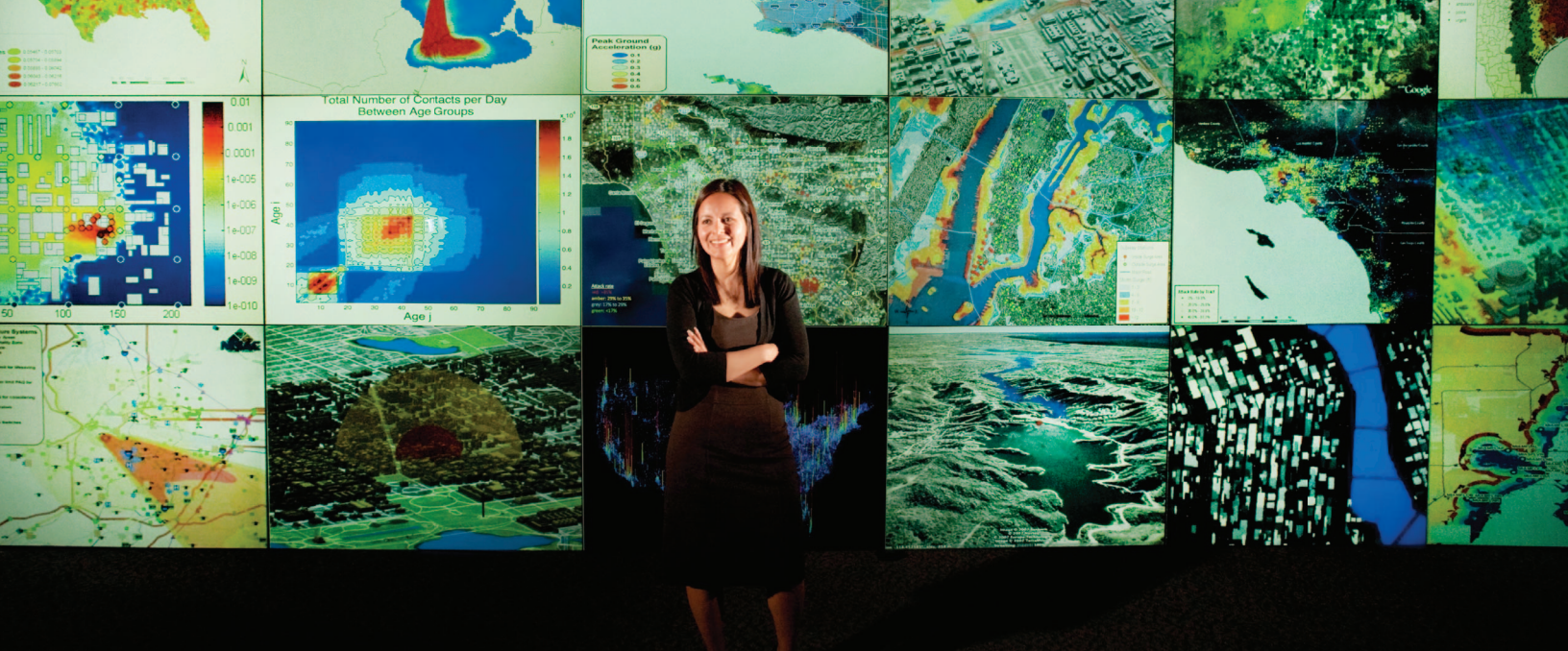
Jason Gans, Murray Wolinsky, Norman Doggett, and Jian Song engage in conversation at their cluster—a collection of computer servers that work together. The team uses this cluster, along with their specialized algorithms, to identify unique signature sections of DNA or RNA that can be used to detect pathogens.

showed high utility for both methods. In some cases, they found that a data stream might only be optimal for a particular disease in a particular country.

“With our massive data streams, we have found that diversity is key,” says Deshpande. “One perfect data stream does not exist.”

This effort laid the foundation for a collaboration straddling military and civilian health surveillance, and the Los Alamos team’s evaluation framework is being considered for disease surveillance as well as other public health initiatives. In addition, the Los Alamos team developed the Biosurveillance Resource Directory (BRD), a relational database that underwent pilot testing by members of the human, plant, and animal disease-surveillance community. The BRD is intended to be a global resource to facilitate rapid information access.





Sara Del Valle stands in front of the Los Alamos powerwall, where multiple computer simulations display the impact of different diseases and disasters across multiple regions.

Path of the storm

High performance computing at Los Alamos has made possible the development of predictive models to help inform decision makers. Once critical details about a disease outbreak are known, a model can be used to forecast how the epidemic may progress and analyze the effectiveness of proposed countermeasures.

The Epidemic Simulation System (EpiSimS) is one of the tools developed at Los Alamos to model epidemics. This model uses several data sources, including U.S. Census data, to create a detailed virtual world in which synthetic people interact and spread disease in a realistic fashion. They go to school, work, and perhaps the grocery store, and they might ride trains or buses at some point during the day before returning home to their families. When an infected individual is introduced, the model can simulate how fast the disease will spread based on the interactions each person has—person A goes to work and visits person B, then person C, then goes to a store and interacts with person D, etc. By incorporating detailed mixing and activity patterns, EpiSimS can estimate which groups of people will be affected and where. This information helps scientists develop targeted mitigation strategies.

A similar system called EpiCast, an epidemiological forecast, was also developed to model epidemics, only faster and with less detail than EpiSimS. For instance, instead of simulating each person's daily interactions, EpiCast uses an average based on empirical surveys and previous models—there are *X* individuals on a given day at home, *Y* in the workplace, *Z* out shopping, and so on. Data describing how many people commute from one census tract (a roughly 5000-person subdivision of a county) to another captures

detailed workflow patterns, and long-distance travel data is used to model less regular mobility.

“This allowed scientists at Los Alamos to do a national simulation of flu season in a few hours, whereas it might take EpiSimS a few hours to do a more detailed simulation of just California,” says computational scientist Tim Germann. Both simulations are fairly accurate; they have been validated against historical outbreaks as well as actual recent outbreaks that Los Alamos has been called upon to examine. In 2006, for example, EpiSimS was used to inform the Department of Homeland Security about preparedness for a potential avian flu outbreak. And in 2009, both EpiSimS and EpiCast were used to forecast the spread of swine flu. In both studies, Los Alamos teams investigated how quickly the disease might propagate, as well as how effective various countermeasures would be.

Los Alamos's logical next step, a multi-scale epidemiology model (MuSE) that incorporates multiple host organisms, couples larger-scale interactions—counties instead of subdivisions—with the small-scale dynamics of disease spread. MuSE was designed specifically for biosurveillance and has been used in recent years to study rinderpest and foot-and-mouth disease in livestock in the United States, avian flu in Nigeria, and Rift Valley fever in East Africa.

Lab scientists are also making use of data from social media, such as Twitter, to inform their models about how people's behavior might foretell the spread of disease. They discovered that people tend to tweet all sorts of details about their lives, including when they wash their hands and whether or not they have been wearing a facemask. This could be a valuable way to track how the public responds to health warnings or recommendations.

“We tried to find if facemask usage correlated with disease spread,” says Sara Del Valle, a computational epidemiologist at Los Alamos. “As the incidence of the disease and the public perception of its incidence go up, people wearing and talking about facemasks go up, and as the incidence declines, so do the usage and mentions. This is crucial for understanding and modeling infectious diseases because changes in people’s behavior can affect the spread of an epidemic by reducing their risk of infection.”

Integrated response

Looking to the future, scientists have been considering how to further expedite disease response by integrating existing databases, analytics platforms, and modeling programs to rapidly evaluate a situation and recommend countermeasures. One example of this kind of integration is a Los Alamos-led pilot project called BioPASS (pathogen analysis supporting system) that demonstrates how existing biosurveillance systems could be accessed and integrated through a user-friendly Web interface. Upon receipt of information from a rapid diagnostic test, BioPASS can access existing genomic databases and analytics to help identify the pathogen in question and create a simple model, showing both how the disease could progress and how the impact could be reduced by certain countermeasures. For instance, it could display a graphical comparison of the effect of administering antibiotics to the patient on day two versus another showing antibiotics beginning on day six.

“The idea is to enable analysis and collaboration using many existing platforms and a variety of data sources,” says Los Alamos biologist Helen Cui. “This will help inform decisions that must be made quickly.” Furthermore, the hope is to take this analysis one step farther—perhaps the incident location could be cross-referenced with Twitter data to identify if there are outbreaks in nearby geographic locations. The BioPASS pilot was very successful, and the team is now proposing to broaden its scope to include more data streams and more variables.

Integrating the components of biosurveillance is a major endeavor, but the scientific community at Los Alamos has been working toward this goal for some time. After the release of the *National Strategy for Biosurveillance*, Basil Swanson served on the review team for the Strategy’s implementation plan. Helen Cui was also a participant in the *Biosurveillance Science and Technology Roadmap* for the Strategy. Through this participation, Swanson and Cui were able to bring a perspective of technological advances to the national biosurveillance picture.

“In general, some of the biggest gaps in biosurveillance are in diagnostics, big data analysis, and modeling. These are all things Los Alamos does well,” says Swanson.

LDRD

—Rebecca E. McDonald

Global Genomics


At the end of the Cold War, the U.S. government worried that nuclear material or biological agents in the former Soviet Union could fall into the wrong hands. To reduce the risk of theft, the United States implemented a number of strategies to help lock up old facilities and secure their contents. The program was called Cooperative Threat Reduction (CTR) and was part of the Nunn-Lugar Act of 1991.

In 2009, CTR expanded to enhance global capabilities for detection and diagnostics and to create a cooperative network that includes a wide range of countries, international organizations, and non-government partners. The purpose was to prevent, reduce, mitigate, and eliminate common threats to national security and global stability.

Los Alamos scientists, drawing upon their success with the Human Genome Project, began multiple efforts in 2012 to share their extensive genomics expertise as part of the CTR effort. One such initiative involves a collaboration between Los Alamos and the Johns Hopkins University Applied Physics Laboratory to create a library of diagnostic assays for especially dangerous pathogens to aid in rapid detection of disease.

In another CTR initiative, scientists from Los Alamos have been instrumental in helping establish sustainable genome science programs at the newly formed Center for Public Health Research (CPHR) in Tbilisi, in the Republic of Georgia, and, in collaboration with the Sandia National Laboratories, the Center for Genomics Science at the Jordan University of Science and Technology (JUST) in Irbid, Jordan. Los Alamos has already delivered sequencing equipment and computer systems and conducted in-house training in Tbilisi as well as helped with the plans for a major renovation in Irbid to house the JUST Center. Scientists from both facilities have also traveled to Los Alamos to receive further informatics training, including a recent workshop in June 2013 that included CTR partners from South Africa as well.

“The Los Alamos CTR program is continuing to expand,” says Chris Detter of the Los Alamos Emerging Threats Program Office. “We will continue to support these international partners by providing specific protocols for sequencing, assembling, and analyzing genetic data. We are also building capabilities in these and other countries as a foundation for a wide variety of research, including pathogen identification and the development of therapeutics and vaccines.”



This visualization of sequence analysis and annotation helps when comparing organisms’ genetic blueprints.