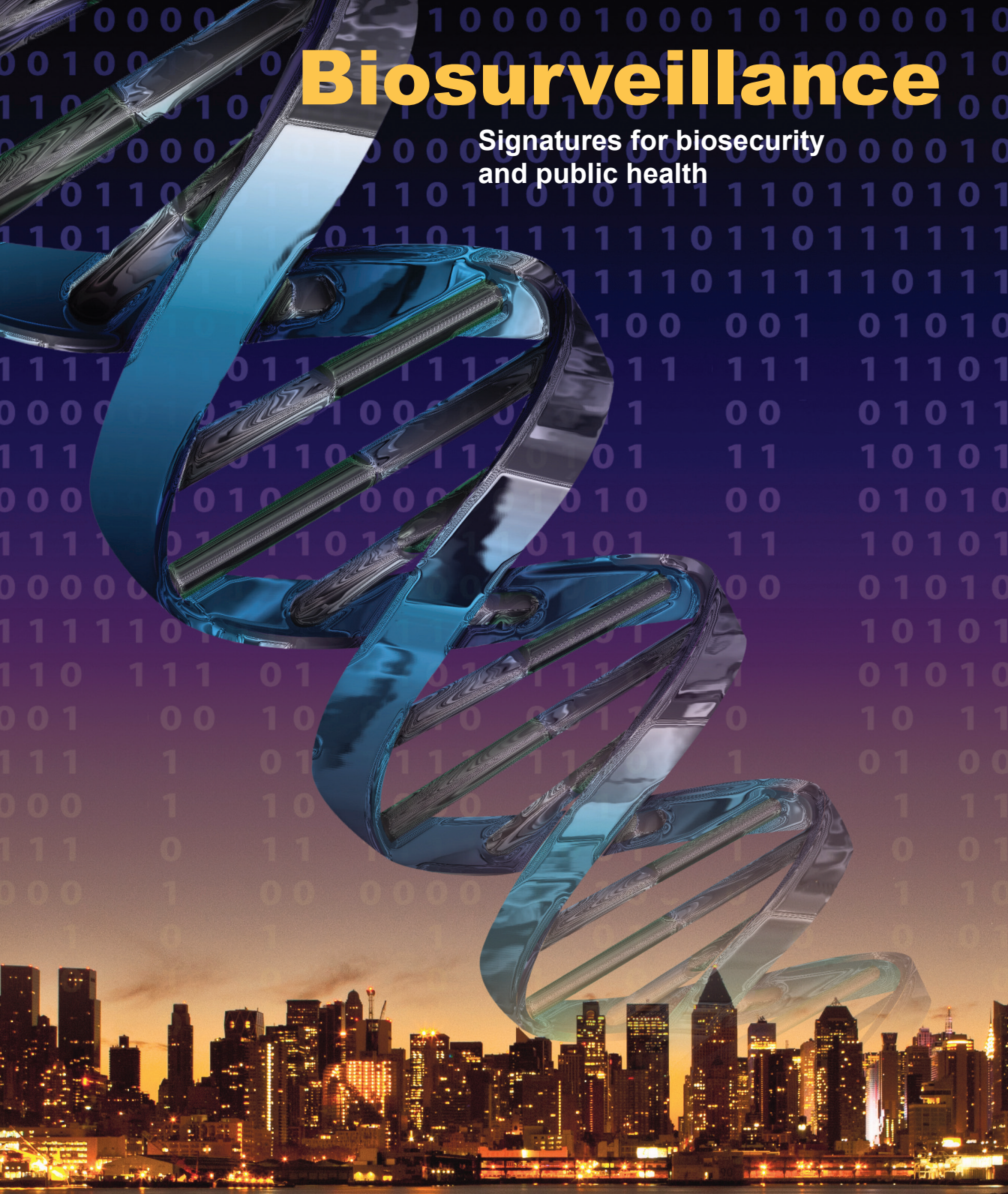
Biosurveillance

Signatures for biosecurity and public health



A Los Alamos National Laboratory
Implementation Plan for the Science of Signatures

Los Alamos
NATIONAL LABORATORY

Biosurveillance at Los Alamos

Los Alamos National Laboratory's charge is to develop science and technology that will make the nation safer and enhance our global standing. This breadth of mission scope requires careful internal planning and effective cooperation with external partners and other governmental agencies. The document you are holding is one of the products of ongoing planning efforts that are designed to bring to bear the Laboratories unique capabilities on problems of the greatest significance.

To those unfamiliar with the extent of our science, it may seem odd that our planning includes such a strong biology focus, yet our work in this area extends all the way back to the Manhattan Project and the birth of large scale government-sponsored research. Following World War II, the Laboratory began programs in health and radiation physics that expanded to become the robust bioscience program we have today. Along the way, we have made contributions such as the Human Genome Project, the invention of flow cytometry, HIV virus management, and more. Our greatest advances come at the interface of biology and high performance computing, modeling, the physical sciences, and engineering.

Our objective now is to look to the future and apply our strengths in these areas to new problems in biosurveillance. The challenge is not small; we must manage and marshal our internal resources, select the correct problems to which those resources should be applied, and then coordinate with external collaborators in academia, government, and industry.

We intend this high-level plan to help us both communicate internally and initiate conversations with the Departments of Health and Human Services, Defense, Homeland Security, and others into how we might apply our expertise to helping solve biosecurity problems.

Additional resources can be requested from the contacts listed on page 18. Strategic planning is a living process, and we welcome your feedback and questions.

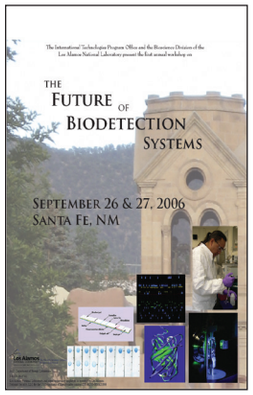
Nancy Sauer is the Associate Director for Chemistry, Life, and Earth Sciences, which oversees the Science of Signatures effort for the Laboratory. She can be reached at (505) 606-2266.

Biosurveillance at Los Alamos is jointly managed by the Science of Signatures and the Information Science and Technology (IS&T) pillars (see p6). Capabilities in supercomputing, large scale data analysis, and modeling combine in a unique way with our more conventional bioscience capabilities to address issues in disease surveillance. Shown here are the supercomputers Moonlight and Cielo. The Cielo system is a classified supercomputer capable of analyzing the most sensitive challenges in biosecurity.

Table of Contents

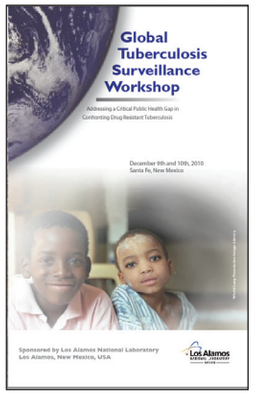
LANL conferences in biosurveillance.

2006



2007

2010



2011

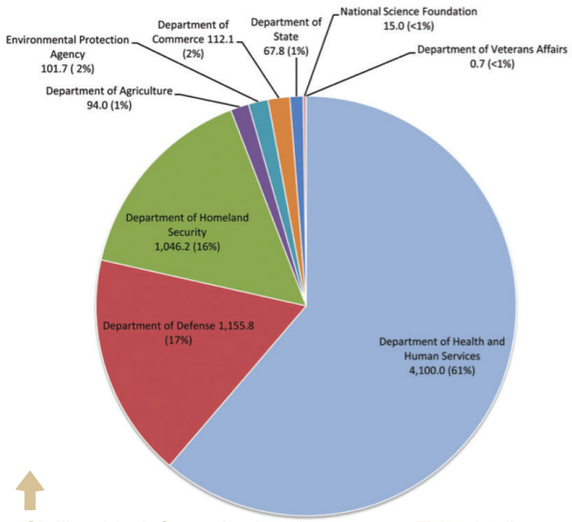
The Laboratory has been involved in organizing national and international biosurveillance technology advancement conferences since 2006.

A National Imperative

There is a recognized national need for advanced biological signature technologies that could, among other things, address disease surveillance. The Laboratory's strategy in biological signatures maps directly to the national posture in biodefense. In 2014, the government is projected to spend almost $6.7 billion on biodefense, with the bulk of that targeted to programs having application in nondefense space as well. See figure, right.

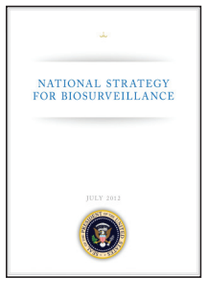
The intent of this funding, which has grown each year from 2012 to 2014, is preparedness, prevention, and mitigation of biological threats to civilians. Each of these areas has a significant, even primary, signature related aspect.

The goal of the Laboratory is to partner with the agencies that are executing our national biodefense strategy and provide science and technology solutions to biological threats using the unique resources of the Laboratory.



Civilian biodefense budget by agency, FY2014 (in $millions). Federal Agency Biodefense Funding, FY2013-FY2014. Tara Kirk Sell, Matthew Watson. Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science. September 2013, 11(3): 196-216.

The Laboratory's approach is informed and guided by both the recently-published National Strategy for Biosurveillance (White House), and a study by the National Academy of Sciences that focuses on the importance of cooperative engagement amongst the nation's science and technology leaders.



The Laboratory provides technology, expertise, and on-the-ground support to the agencies tasked with managing biosurveillance threats. We do so in collaboration with a wide variety of partners and collaborators.

Laboratory Planning

Strategic planning necessarily cascades through an organization, and this biosurveillance overview document grew out of a lengthy planning process that involved both internal and external workshops and reviews. See the organizational chart in the next section for a sampling of the organizations that contributed.

Beyond this biosurveillance planning, the Laboratory is engaged in numerous other strategic exercises. Collectively their intent is to coordinate and marshal people, equipment, and facilities behind important technological problems within the scope of our mission.

Who Should Read This

This document was written with multiple audiences in mind.

Our staff: This document is the high level Laboratory strategy for biosurveillance. It is intended to unite Laboratory scientists behind a single strategic direction and to the guide the research and efforts of our staff. It will inform investments in hiring, facilities, equipment, as well as additional planning.

Our stakeholders: It is intended as a communications document for those outside the Laboratory who have current or potential programs that might benefit from the expertise available at Los Alamos. With it, we hope to begin a dialog that will help us understand the technological and scientific requirements of our customers and to inform them of where we might be of assistance.

Our collaborators: Biosurveillance is a cross-organizational and multidisciplinary endeavor. To be successful, we must collaborate with other science and technology organizations, the public health sector, local and national governmental bodies, and through them with the civilian population that we seek ultimately to serve. This document explains what we hope to achieve and where we will require partners to succeed.

The National Security Sciences building from the Los Alamos townsite.

LANL Strategic Context

The science pillar concept is a primary tool the Laboratory uses to plan how we will accomplish current and future missions, including the biosurveillance mission that is the focus of this document. There are four science pillars: 1-Materials for the Future (Materials), 2- Integrating Information Science and Technology for Prediction (IS&T), 3-Nuclear and Particle Futures (NPF), and 4-the Science of Signatures (SoS). Each of the pillars has discrete science goals that are fundamental to building the Laboratory's future science and technology base. These pillars support each other, and interfaces among the four pillars are leveraged for the benefit of all four.

The fundamental precept of this approach is that the greatest science breakthroughs will come as we approach difficult problems in revolutionary ways. This multidisciplinary approach draws upon physicists, materials scientists, chemists, computer scientists, theoreticians, biologists, earth scientists, space scientists, engineers, mathematicians, and numerous other disciplines to solve important national security science problems. The pillars approach gives these experts a framework for working together and allows them to apply their skills across the traditional boundaries of their disciplines.

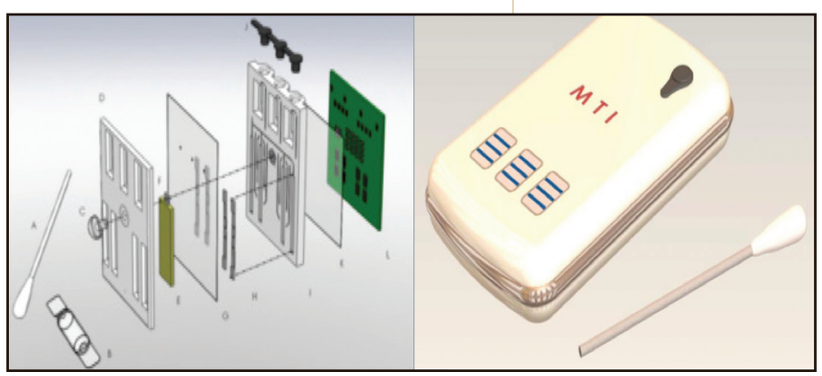
The science pillars also inform our investments in science and engineering, guide recruitment and training strategies, and serve as a framework for our partnerships with other leading research institutions worldwide. Effective biosurveillance incorporates a wide variety of signatures and thus ownership of the associated strategic planning efforts fall under the Science of Signatures. However, there is a significant component of both IS&T and Materials required in biosurveillance. Laboratory planning efforts in biosurveillance have therefore included leadership from these pillars as well.

Science of Signatures

The Science of Signatures strategic plan (published 2012) was chartered by the Principal Associate Directorate for Science, Technology, and Engineering (PADSTE) and written by a team of more than 35 managers and scientists from across the Laboratory following a two year process. Its grounding principle is that we must be able to identify and characterize threats before we can understand them or take action to mitigate them. Signatures are the unique elements that allow us to locate threats within their environments and describe them.

The Science of Signatures pillar addresses emerging challenges by developing science and technology to detect these threats. Our complete technological toolbox is applied to signature science from across our mission areas of global security; nuclear defense; and energy, climate, and health. Critical components are the discovery and detection of signatures to enable understanding of component species or processes that have a major impact on a large, complex system.

Specifically, we characterize measures, signals, and properties in or of complex systems to order to detect or attribute change; predict systems behavior across scales in space (molecular to global) and time (near-term to geologic), and assess impacts to the system of change.



From concept to application: this hand-held "dipstick" was developed by LANL scientists to detect multiple pathogens quickly and with minimal resources. The scientists who invented it are now pursuing its development through the startup company Mesa Tech International.

SoS Scientific Approach

Our scientific strategy is to *discover new signatures*, *revolutionize measurement of signatures,* and *deploy new technologies in the field*. Each of these three components has distinct characteristics.

Discover signatures: Identify signatures of chemical, biological, radiological, nuclear, and explosives threats and of climate, energy and health security impacts. In essence, signature discovery is determining those measurable phenomena that uniquely identify and characterize properties within complex environments.

Revolutionize measurements: For threat-specific signatures, develop entirely new measurement technologies, methodologies, or strategies or develop transformational advances in the current state-of-the-art. In essence, how can sensitive and specific measurements be made in entirely new ways and/or how can new phenomena (signatures) be measured?

Forward technology deployment: Make measurement technologies and methodologies practical through engineering. This includes prototyping of sensors and instruments for field deployment and systems integration of sensor networks. In essence, how do we bring science advances to the real world in a way that provides feedback into signature discovery and/or revolutionary measurement technologies?

The strategy of *Discover, Revolutionize, Deploy* is applied to the Los Alamos SoS areas of scientific leadership:

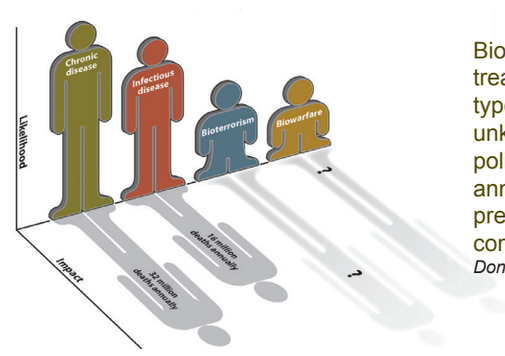
- **Nuclear and Radiological**
- **Chemical and Materials**
- **Biological**
- **Energy**
- **Climate**
- **Space**

The other five areas of leadership are developing strategic plans similar to this biosurveillance plan. Each individual plan will nest with the larger SoS plan.

SoS is tightly integrated with the other pillars along the themes of Discover, Revolutionize, Deploy.

The Science of Signatures Pillar is

- An Institutional effort led by the Associate Directorate for Chemistry, Life, and Earth Sciences.
- An organizing principle built on historical strengths and technical leadership.
- One of four science plans to shape and manage the future of science and technology at the Laboratory.
- Supported and mandated by DOE.
- Important to our immediate stakeholders and the nation.
- Reshaping the way we think of multidisciplinary science.



Biological approaches to detecting pathogens, diagnosing and treating disease, and predicting spread are useful for mitigating all types of biothreats. The impacts of bioterrorism and pandemics are unknown but can potentially be great when considering the social, political and economic effects that inevitably follow. By comparison, annual deaths from chronic and infectious disease are comparatively predictable in both impact and likelihood but are still very destructive. A comprehensive plan for biothreat reduction must consider both.
Donald Montoya, graphic.

Biosurveillance Overview

Biosurveillance is the collection, analysis and interpretation of data to help monitor for he presence of pathogens in plants, animals, humans, food, and the environment. Its goal is to save lives by informing intervention strategies and by guiding public health decisions. Biosurveillance is the central theme of biosecurity.

Our vision is for the Laboratory to integrate and use unique strengths in biological, physical, and information sciences to safeguard the nation through advances in biosurveillance.

Historically, biosurveillance has been dominated by the military's concern for threat pathogens (commonly called "select agents"), but more recently its scope has broadened to include identification of emerging infectious disease and a concern for the civilian populations. An additional permutation is a concern for safeguarding US military personnel in regions where infectious disease coexists with diseases caused by threat pathogens. The ultimate goal of the evolving national effort in biosurveillance is an intelligent, real-time system where differing surveillance data streams, which range from social media to clinical and point-of-care diagnostics, can be integrated and rigorously analyzed to provide guidance to decision makers from the local to the national level.

To succeed, the national effort will require new tools and approaches for

- **Point-of-care diagnostics,**
- **Predictive modeling, and**
- **Decision support through complex data analysis.**

These three goals (which are detailed in the following pages) play to the Laboratory's current and historical strengths and thus comprise the structure of our biosurveillance strategic plan. An integrated approach allows the Laboratory to draw upon the multidisciplinary science that has been our hallmark for more than 70 years.

For instance, Los Alamos experimental biologists specialize in novel *in vitro* methods to generate reproducible, high-quality antibodies and other affinity reagents such as peptides. This approach generates better results and avoids the use of animals in research. These capabilities, when combined with other LANL expertise in biomarker discovery and assay development, have extensive applications for detection and diagnostics. Other strategies for diagnostics include nucleic acid-based signatures and genomic sequencing, also areas in which Los Alamos has made significant contributions.

Theoretical biologists and epidemiologists at Los Alamos are widely recognized for their development of agent-based, predictive models that can reveal the potential path of an outbreak, as well as the effect of different mitigation strategies. This work leverages a strong Laboratory high-performance computing program that includes some of the fastest super computers in the world.

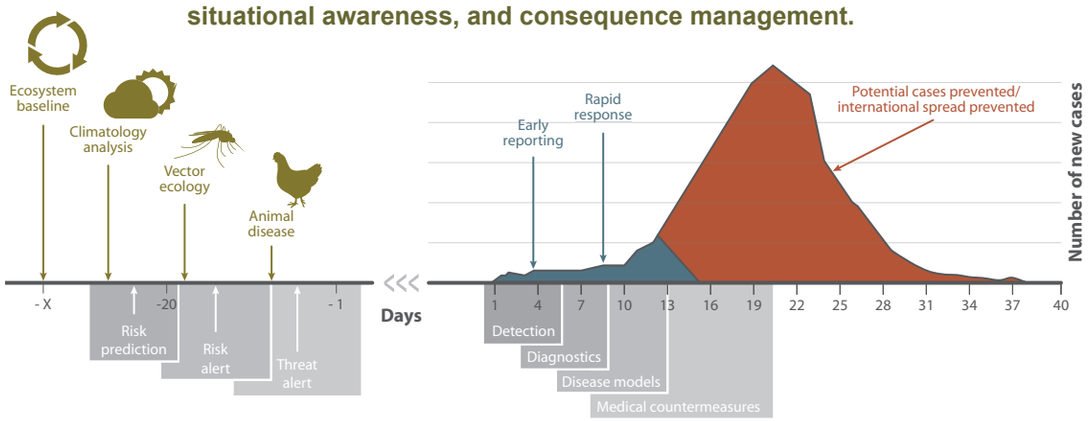
Laboratory capabilities in Information Science and Technology (e.g., computational co-design, data-science-at-scale, analysis of complex networks) have also been used to address "big data" problems that range from climate modeling to cosmology and are characterized by the integration of diverse and heterogeneous data streams with quantified uncertainties.

These strategies are now being used to address the formidable data integration and analysis problem required by an intelligent, real-time biosurveillance system. Additionally, other LANL capabilities not previously used in biological research show potential for our evolving biosurveillance effort

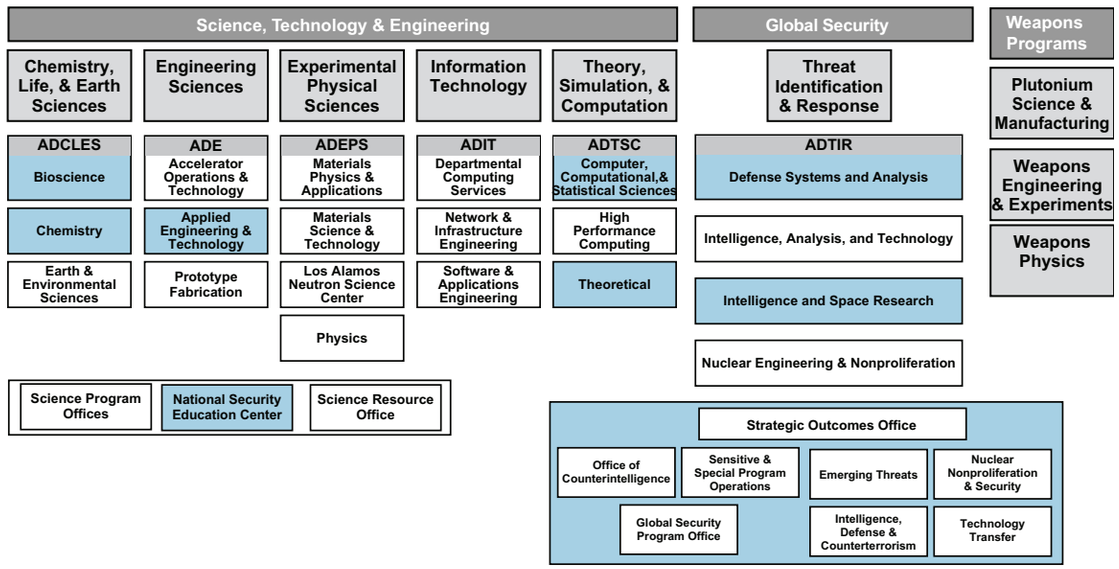
Los Alamos has a rich legacy of engineering capabilities that have traditionally been used to build spaced-based instrumentation and sensors as well as remote and unattended sensors. These capabilities have in the past and could in the future be used to adapt and deploy detection and sensing discoveries made on the bench to real-world applications in the field and in clinics.

Finally, image analysis and machine learning capabilities that have been developed at the Laboratory to extract knowledge from high content but poor resolution images. High volume sensing data are being explored as a means to extract diagnostic information from image-based sensors. These emergent capabilities, as well as those already being used to address biosurveillance, form the bedrock of our approach. As shown in the organizational chart below, this strategic plan requires that expertise be drawn from across the Laboratory.

Biosurveillance components: early warning, early detection, situational awareness, and consequence management.



This timeline of events in a hypothetical disease outbreak shows that detection and diagnosis in the first week can enable a rapid response that would significantly reduce the overall spread of the disease. However, looking backwards in time, 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. A comprehensive biosurveillance strategy spans the spectrum. *Leslie Sandoval, graphic.*



Laboratory organizational chart showing in blue the organizations that currently contribute to biosurveillance science development. The Laboratory as a whole has a budget of over $2B and more than 9000 employees, and while our "bench strength" is already great, it can clearly go much deeper should the science challenges demand it.

Tuberculosis Outbreak

Case Study

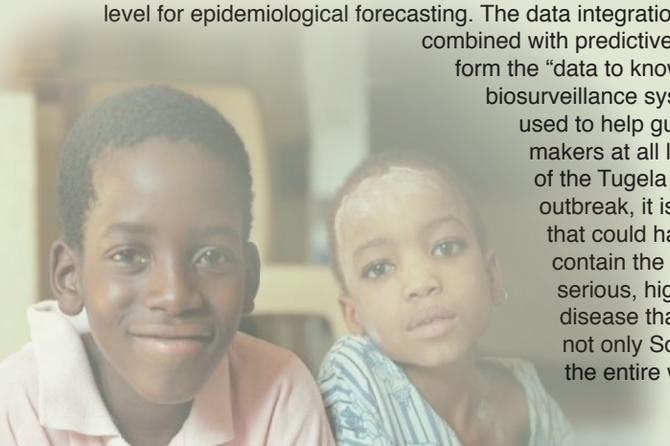
In 2004 in the small village of Tugela Ferry in Kwazulu Natal, South Africa, 54 people tested positive for *Mycobacterium tuberculosis* (TB). In 16 days, 53 had died. The rapid progression to death in this population caught the world by surprise. TB is generally thought to be a chronic, slow moving bacterial infection that is readily treatable with a cure rate of nearly 100%. But the Tugela Ferry strain was different, partly because the infected individuals were HIV positive, but mostly because this was a new mutant strain of TB. It was resistant to not only first line drugs but also second line drugs developed recently and used to treat the multi-drug resistant strains (MDR/TB) that began appearing in the 1990s.

In less than a decade, this extensively drug resistant TB (XDR/TB) has spread throughout South Africa and in recent years has been found in many other countries. One of the world's leading experts on this outbreak (Willem Sturm of the University of Kwazulu Natal and then Director of the Kwazulu Natal Research Institute for Tuberculosis and HIV) believes the TB outbreak could have been stopped at Tugela Ferry with a comprehensive biosurveillance strategy similar to what is proposed.

Diseases such as this will continue to emerge and we must prepare. An effective and comprehensive biosurveillance strategy has a flexible approach that can be implemented in different scenarios: identify and stop the disease at inception; contain its spread; inform local, national, and global strategies for management. In Tugela Ferry, better diagnostics would have made a big difference; authorities needed to know who had active infection. Most current TB tests are fraught with issues—some can show false positives in those with a prior infection or a vaccination, while others can show false negatives in certain populations (those with HIV, for example). Moreover the tests must be feasible in resource poor settings such as Tugela Ferry.

The rapid and accurate field identification of active infection allows immediate initiation of treatment and quarantine protocols to minimize further spread. It also provides a filter to triage patients for more lab-intensive strain identification. In a few hours, laboratory methods such as PCR can identify strains which in turn generates a recommended treatment protocol, indicates prevalence of drug resistant infections, and describes transmissibility. Sequencing technologies provide the next level of characterization and facilitate the identification of new strains and previously unreported mechanisms of drug-resistance.

This diagnostic approach provides a reliable data stream for inclusion with many other discordant data streams (e.g., syndromic and social media data) that can then be used to track disease progression at the population level for epidemiological forecasting. The data integration and fusion combined with predictive modeling can form the "data to knowledge" biosurveillance system that can be used to help guide decision makers at all levels. In the case of the Tugela Ferry XDR/TB outbreak, it is this information that could have been used to contain the spread of this serious, high mortality disease that now threatens not only South Africa but also the entire world.

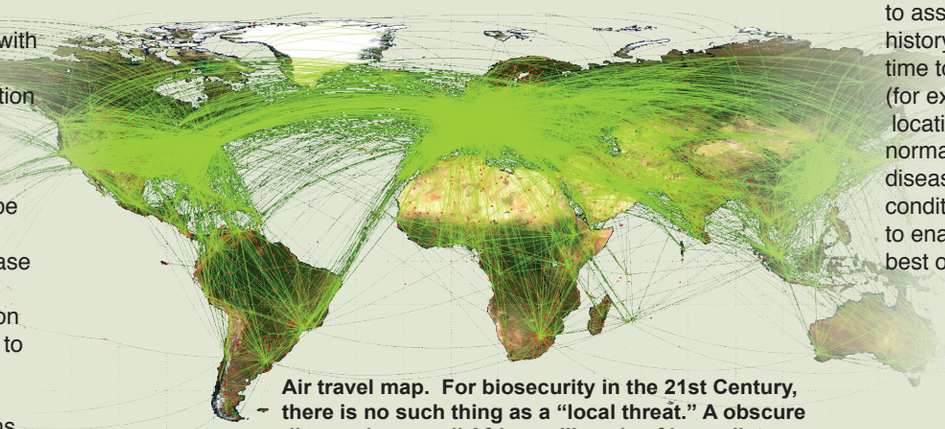


A Path Forward: Los Alamos National Laboratory can serve as a science and knowledge resource for the agencies and entities that are the first line of response. The three part strategy outlined in this brochure can be a major force in mitigating biological threats such as the XDR/TB outbreak, a bird flu outbreak, or an emerging disease that makes a species jump.

Prompt diagnostics: Multiple efforts at LANL are improving detection through field-ready diagnostic tools such as the optical biosensor that could quickly and more reliably detect active infections: e.g., XDR/TB or H1N1.

Data integration and analysis: Integration of disparate data sources for analysis, including social media such as Twitter, can provide valuable information on emerging disease hot spots in real time.

Decision support through complex data analysis: Epidemiological models at LANL have been used in past epidemics to predict disease progression and to help identify measures that could be taken to stop it. Other models are used to help identify potential outbreaks using animal disease models and climate data.



Air travel map. For biosecurity in the 21st Century, there is no such thing as a "local threat." A obscure disease in a small African village is of immediate importance to the US, as is the development of new biothreat agents.

WikiCommons

Bird Flu Outbreak

Case Study

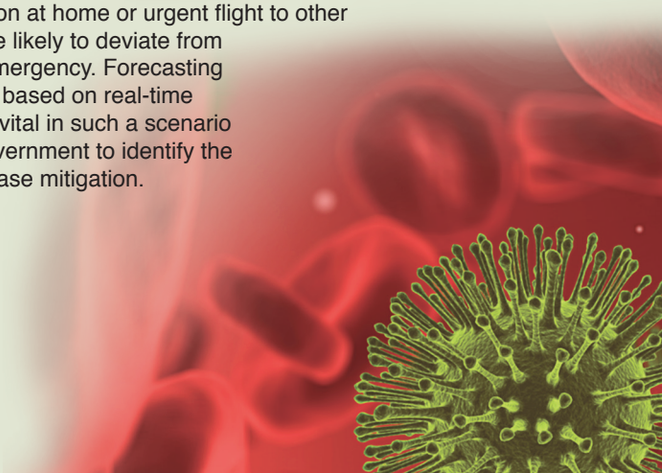
In April, 2009, a 10-year-old boy in San Diego, California, was diagnosed with a novel strain of H1N1 flu. Newspaper headlines asked if it was a bioterror attack. The query arose again in 2011, when a highly unusual strain of *E. coli* spread rapidly through Germany.

Every time a novel pathogen appears—on average, one every year—the "what if" question arises. ***What if*** that flu strain had instead been a deadly H5N1 avian flu engineered by terrorists? Release of a high-fatality H5N1 strain able to spread rapidly among humans could result in millions of deaths. Clinical diagnosis of a terrorist attack would be too late for optimum containment. The most effective way to reduce biological threats to national security is to avert them early. A major focus of LANL science is the identification of threat signatures and technology development for early detection—before anyone gets sick.

In 2012, the international community panicked when US and Dutch scientists evolved potentially human-transmissible strains of H5N1 bird flu, which has a 60% mortality rate in humans. Concern focused on the potential for terrorists to misuse the research as a recipe for a bioweapon. Los Alamos is developing capabilities to counter this type of emerging threat by integrating information and technologies from a variety of sources. They seek to disrupt early phases of threat development in various ways, including enhanced export control, which attempts to reduce the spread of enabling technology. LANL is also developing capabilities to enhance detection of technical expertise recruitment and improving tools to detect the acquisition of reagents and equipment for nefarious biological laboratories.

If terrorists were to succeed in producing a human transmissible strain, poor containment during development could result in a local outbreak of highly-fatal influenza, which would produce threat signatures for detection Poor containment during development could result in a local outbreak of highly-fatal influenza, which could produce other threat signatures for detection. To improve early detection of anomalous events like these, LANL is contributing to efforts led by DoD, the State Department, and DOE to enhance disease surveillance capabilities in central Asia, the Middle East, and Africa.

Finally, if the enhanced strain was released, CDC and DHS would recruit a set of epidemiologists to forecast the progression of the epidemic in order to assess the possibilities for containment. LANL continues to build on its history of producing cutting-edge disease forecasting by developing real-time tools that will incorporate human behavior patterns (for example seclusion at home or urgent flight to other locations), which are likely to deviate from normal in such an emergency. Forecasting disease progression based on real-time conditions would be vital in such a scenario to enable the US government to identify the best options for disease mitigation.

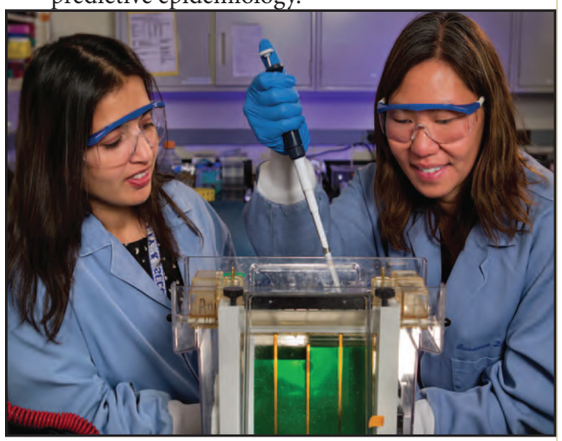


Goal 1: Diagnostics

Early detection of biothreats is critical for controlling their spread. Clues for an impending outbreak are varied and require creative, orthogonal approaches such as monitoring animal populations and climate data to find emerging pathogens, as well as communication data to intercept terrorist plots.

Timely detection of disease that is already active in the human population is often constrained by the limits of diagnostic tests. There is a broadly acknowledged need for better point-of-care (POC) diagnostics. However, to realize a future goal of reliable POC diagnostics for all infectious diseases, known and unknown, it is critical to develop a new toolset for pathogen identification and characterization.

LANL's technology roadmap for biosurveillance includes this critical requirement for thorough characterization of pathogens at both regional and advanced research laboratories. Pathogen characterization is critical to the identification of new strains and serovars, predicting pathogen evolution and transmissibility, and tracking emerging drug resistance. All of these factors are critical for determining the path forward in the design of vaccines and countermeasures, and are also significant for real-time biosurveillance through predictive epidemiology.



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.

Drug resistance is increasingly a problem in both bacterial and viral pathogens, and often the mechanism of resistance is unknown. Understanding anti-microbial resistance is critical to effective countermeasures. Further, patients in regions like East Africa manifest with multiple health issues which can complicate diagnosis: co-infection with other pathogens (e.g., tuberculosis), species jumps from wild animals to humans, vector-borne diseases, simultaneous infection from multiple pathogens, and myriad chronic infections and existing co-morbidities (HIV, malnutrition, diabetes, and others).

Understanding of the above factors is critical to achieving real-time global biosurveillance, which integrates effective forecasting of outbreaks using diagnostics and modeling.

Accordingly, the LANL science community is focused on diagnostics for effective biosurveillance. These encompass field-ready diagnostics for the identification of active infection (for uses as varied as soldiers in remote field locations to home-based diagnostic kits), to more involved/comprehensive assays for strain identification and characterization. There are at least three critical technological approaches that can help us realize this goal:

1) Fieldable/rapid diagnostics for immediate pathogen identification,

2) Nucleic-acid-based assays for identification of known strains and markers of drug resistance,

3) Genome sequencing (DNA and RNA) for comprehensive identification of all pathogens and novel markers of drug-resistance.

Each of these technical approaches offers complimentary information that together can help in the more efficient identification of the pathogen, minimize redundancy and decrease wasted effort. Together, these technologies allow us to address the following strategic objectives for efficient diagnostics.

Goal 1.1: Create universal approaches for the discovery and detection of all pathogens by

- The discovery of pathogen biomarkers directly from infected hosts.
- Development of new strategies for early detection/diagnosis of infection.
- Development of approaches for discriminative detection: early detection should be accompanied by discriminative diagnosis to facilitate effective countermeasures.
- Innovation for accessory tools and reagents, including recognition ligands, for effective biodetection.

Goal 1.2: Improve strain identification for characterization of drug resistance and related virulence factors by

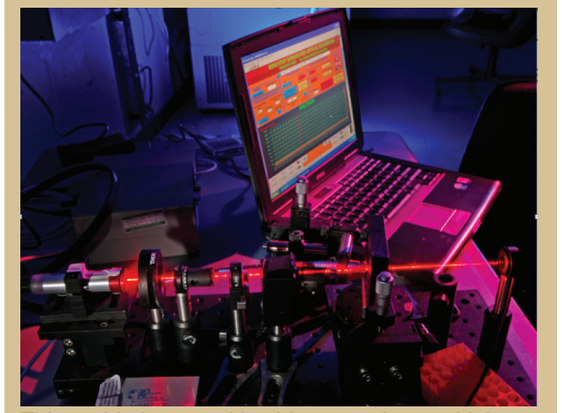
- Advancing phenotypic vs. genotypic detection. Linking phenotypic antibiotic resistance (MIC levels) with genotypic variation will reveal the genetic mechanisms responsible for the various levels of antibiotic resistance.
- Detecting/characterizing known vs. novel antibiotic resistance mechanisms.
- Creating culture-free strategies for genetic analysis.
- Resolving the sequencing challenge vs. informatics challenge.

Goal 1.3: Develop smart sensor interfaces adaptable to any transduction system to facilitate deployment.

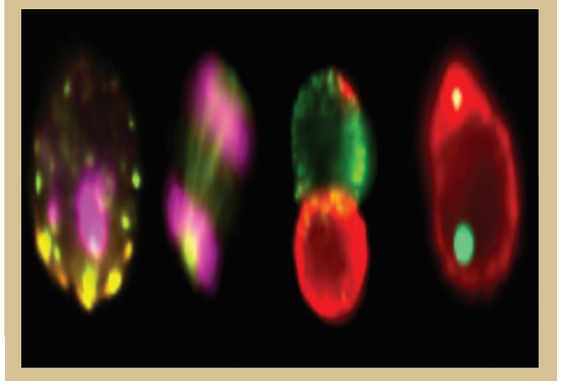
- Detection systems and sensor prototyping for POC diagnostic applications in biosurveillance.
- Data fusion and extraction.

Right: Our ultimate goal is a universal approach that detetcs all pathogens. One possibility is functional validation of sRNAs in virulence (pathogen) or immune response (host). A basic understanding of organism interactions can lead to ways to detect infections, potentially prior to manifestation of symptoms. Shown here is a high content cellular analysis.

Los Alamos computational biologists develop and validate genetic markers for high-consequence pathogens for potential use by DHS BioWatch and CDC Laboratory Response Network labs. Pictured below is a BioWatch detector in front of the Capitol Building in D.C.



This rapid ultra-sensitive biosensor is a multiplex, multi-channel detection system that uses photo-stable and tunable quantum dots. Optimized for field-based and environmental detection, it has been validated for the detection of cholera toxin, influenza, anthrax, tuberculosis, breast cancer, *E.coli*, and other agents using single mode waveguides



Goal 2: Predictive Modeling

LANL scientists integrate predictive models with unique data streams and assays to address biosurveillance needs. These models allow the Lab to collaboratively solve related public health problems in order to both cost-share and greatly improve the fidelity with which the national security mission is carried out.

Predictive modeling capabilities have already been applied to complex health problems (such as methicillin-resistant *Staphylococcus aureus*), anticipating and preparing for a naturally occurring epidemic (e.g., influenza), or assessing the risks from or responding to bioterrorism (e.g., engineered pathogens). Models can fill in incomplete information, enable uncertainty quantification, and guide information collection. They fall into three general categories.

1. Epidemic simulation can project outcomes by using existing infection rates and varying assumptions about transmission mechanisms, behaviors, and control strategies with a goal of informing and optimizing decision-making during an outbreak.

2. Evolutionary analysis predicts how diseases will change. When performed on appropriate sets of pathogen samples it can identify signatures to guide development and use of diagnostics and medical countermeasures.

3. Disease models relating molecular mechanisms to virulence, drug resistance, and immune response can provide important insights when potentially novel pathogens are observed and can guide the development and interpretation of the numerous potential sources of clinical, laboratory, and non-medical information when computing prevalence, risk factors, and efficacy of control measures.

The models developed at LANL can be combined for use in a variety of situations. Effective epidemiological models in biosurveillance require appropriate prevalence data and an understanding of pathogen transmission, risk factors influencing outcome, and quantification of the effectiveness of mitigation strategies. In some instances, such as in the 2002-03 Severe Acute Respiratory Syndrome (SARS) outbreak, analysis of high quality data has allowed effective control of an epidemic.

In other instances, such as the 2009 pandemic influenza, the role of epidemic forecasting can be used to determine the allocation of scarce vaccines and antivirals, identify risk factors of poor outcome, and optimize case definitions for patient care and implementation of disruptive public health measures. In this type of situation, an important attribute of a successful predictive model is its ability to effectively communicate likely scenarios, potential high-consequence scenarios, potential "windows of opportunity" for control measures, and situational awareness of sufficient detail to communicate risk effectively to the public and implement control measures.

Finally, when integrated these three types of predictive models can help anticipate the emergence of novel pathogens—as well as to identify the sample streams from across the globe required to address the problem.

The Lab's capabilities that enable predictive modeling include:

- Agent (subject) based models/discrete event models.
- Analytical models.
- Host response modeling.
- Genomic signatures for host/pathogen characterization.
- Modeling zoonotic reservoirs.
- Uncertainty quantification (UQ).

Based on these capabilities Los Alamos has developed the following strategic objectives in Predictive Modeling for the next five years:

Goal 2.1: Build epidemiological models capable of incorporating disease mechanisms from diverse data sets.

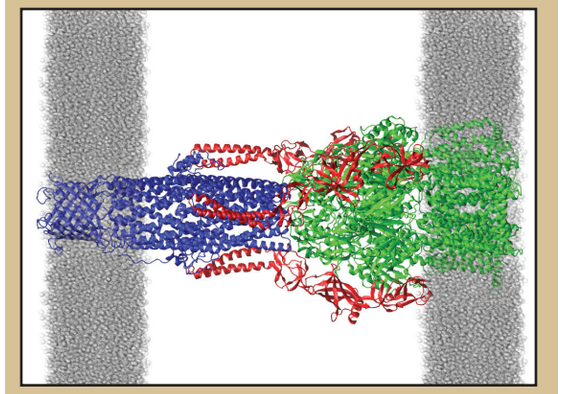
- Assess the risk of *Staphylococcal* and *Salmonella* co-infection in pediatric malaria patients.
- Assess the risk of co-infection from HIV and TB.
- Analyze the impact of vector-borne pathogens in biothreat scenarios.

Goal 2.2: Develop theory and modeling to predict disease progression in a host.

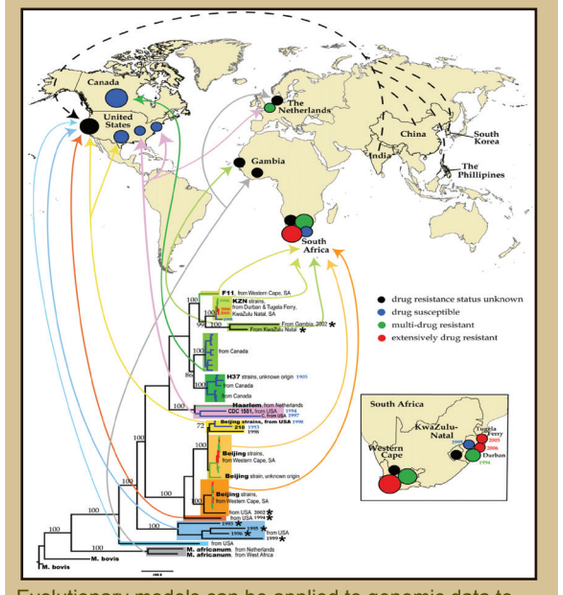
- Develop a link between pathogen levels, target cell death, immune response and disease symptoms in a host.
- Predict how the metabolic state of the target cell is correlated with virus production.
- Model the effect of drug resistance and emergence of a new pathogen in disease progression in a host.

Goal 3.3: Provide decision makers with validated epidemiological models for disease forecast and prediction.

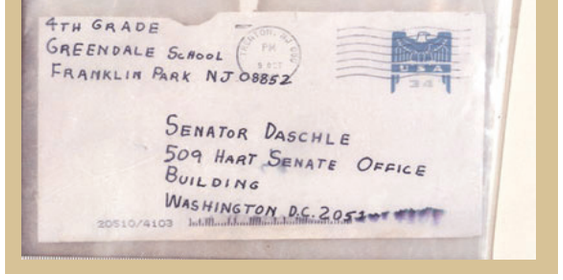
- Benchmarking of epidemiological models.
- Application of validated epidemiological models.



A model of a multi-drug efflux pump in *Burkholderia*; proteins such as this one are a leading cause of multiple-antibiotic resistance. A LANL interdisciplinary team seeks to understand how such genes achieve substrate specificity and how their expression and activity are regulated so that they can be effectively targeted with a new generation of antibiotics.



Evolutionary models can be applied to genomic data to understand global drivers of epidemic spread, identify functional markers, optimize information content from diagnostics, and guide sample collection, as in this example from Ford, et al. published in Tuberculosis in 2012.



Right: Following the "Amerithrax" attack in 2001, Los Alamos scientists began optimizing methods of detecting anthrax and modeling risks. At the time, public health authorities grappled with the idea of a widespread crisis: how many letters would there be, what are the risk factors for various populations, should therapeutics be disseminated or are the repercussions too great? Many of these questions could now be addressed using LANL modeling and detection techniques.

Goal 3: Complex Data Analysis

This goal leverages advances in information science and technology (IS&T) to help public health, agricultural, and national security agencies make timely decisions during outbreak and biothreat scenarios. Doing so requires collection, processing, integration, and analysis of multiple data streams and the development of data analytics tools to transform data into actionable information. This information can then be used by federal agencies, the private sector, international, state, local, tribal governments, and nongovernmental organizations.

Complex Analysis of Big Data: 21st century biosurveillance involves complex analysis of "Big Data," which requires integration, analysis, and interpretation of large disparate data gathered from multiple sources. These sources include clinical diagnosis, syndromic analysis, 'omics studies on clinical samples, biomedical research literature, electronic health records, climate/geographical surveys, lab, imaging, and pharmacy systems, infrastructure systems, on-site and remote patient monitoring, email, and social media. Analysis requires tools and techniques for acquiring, cleaning, and exploiting data using statistical techniques. The results enable situational awareness and provide robust answers to questions of interest.

As in other applications (climate, cosmology etc.) Big Data in biosurveillance has the attributes of four V's: volume, velocity, veracity and variety. "Volume" refers to the size of the data set. High definition spatiotemporal video images or 'omics data can quickly fill up multi-terabyte disks. However, it is not just the sheer data size that must be managed; the "velocity" with which the data is accumulated is also an important factor. It might come from Twitter, Facebook, or sensors.

To be effective, it is necessary to rapidly scan, store, process, and analyze this data. Additionally, many biosurveillance data streams are biased and incomplete and require up-front data assessment and cleaning as well as error quantification to address trustworthiness (or "veracity"). Finally, "Big Data" contains a great "variety." They are collected from multiple sources and capture widely varying data sets. All of these considerations must be taken into account for effective decision support through complex data analysis.

Technical challenges:

- Real data streams are messy, vary in quality, and may be biased or incomplete.
- Relevant data are complex, highly dimensional, and have widely different structures, possibly requiring aggregation and/or disaggregation
- Many relevant data streams must be analyzed locally rather than ported to a central site for analysis by virtue of privacy concerns. This requires computational co-design of appropriate hardware, software, modeling, and data analysis tools that anticipate analyses for local sites (the small) that will scale to all sites (the big).
- Data volume may be large and continuously generated, exceeding the rate (velocity) at which they can be analyzed or stored.
- There are limits to our ability to formulate complex queries and convey the associated uncertainty inherent in the analysis process.

Our multi-disciplinary capabilities address these challenges and achieve efficient data driven decision support through complex data analyses.

Transforming data into actionable information: Situational awareness is important for both early warning and early detection of outbreaks. Analytics and tools that furnish information on how an infectious outbreak would either emerge or unfold enhance situational awareness for decision makers/analysts/public health officials and support planning for prevention or mitigation.

Data sharing and expert analysis of incoming information are equally critical. LANL is developing new capabilities for decision support in infectious disease surveillance. Robust and comprehensive frameworks for data stream and model characterization underlie these tools – they describe fundamental elements of infectious disease surveillance. The resulting data transformation tools are intended to be accessible to the global biosurveillance community. We will build new partnerships and collaborations through subject matter expert panels developed in conjunction with these tools. LANL IS&T capabilities will be leveraged to address the sustainability challenges for developed tools and resources and add functionality.

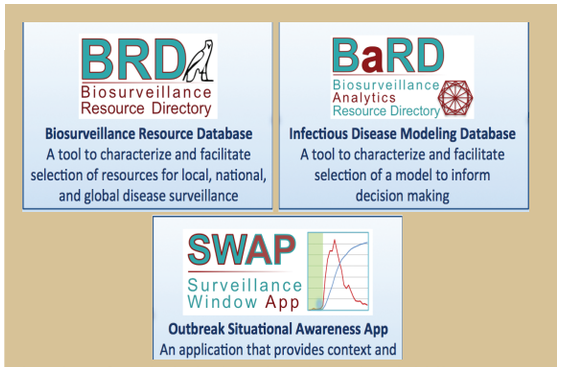
Goal 3.1: Develop and assess new data streams that can be used for disease surveillance.

- Obtain reliable data streams and data sets, and assess content and quality.
- Develop new data streams and characterize them in the context of biosurveillance decision-making. Complex data analysis requires a combination of hypotheses and discovery-driven approaches.
- Develop an evaluation framework to assess both traditional and non-traditional data streams.

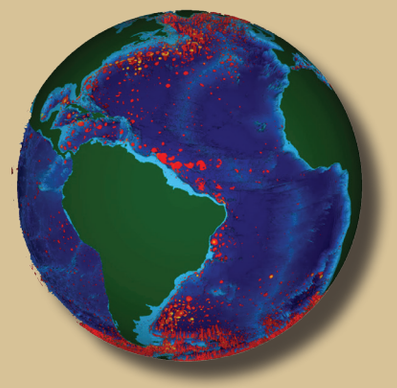
Goal 3.2: Develop and integrate LANL capabilities in disease biology, public health, medical diagnostics, statistical and computational analysis, and computer science to create complex data analysis and decision support tools.

- Identify data streams that can be used synergistically.
- Develop robust integration algorithms for disparate data streams that can be used for disease surveillance.
- Develop robust statistical models and data visualization models that process complex and diverse data streams and support situational awareness.
- Build computational infrastructure to support efficient storage, access, and real-time distributed processing and analysis of biosurveillance data.
- Develop tools and apps that enhance situational awareness in an ongoing event of infectious disease spread.

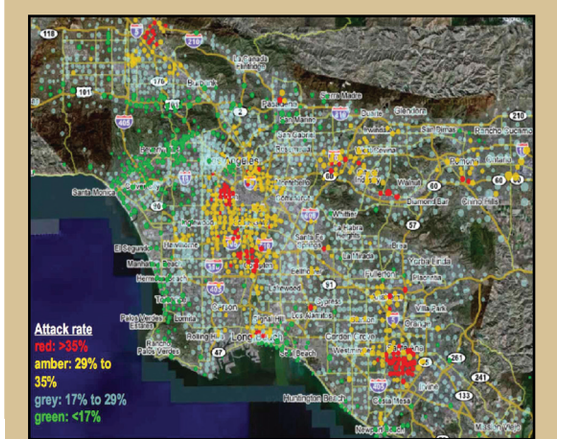
EpiSimS-computed attack rate by census tract for a 1918-like pandemic influenza in southern California. EpiSimS is a stochastic, agent-based simulation that models the spread of disease in regions, allowing for the assessment of disease prevention, intervention, and response strategies. It explicitly represents the daily movements and interactions of synthetic individuals in a city or region, including their interactions with others. It is used as an experimental test bed for analyzing the consequences, feasibility, and effectiveness of response options to disease outbreaks.



A suite of tools is being developed at LANL to enhance situational awareness of an ongoing event and transform data into actionable information. The BRD is a resource for validating information about disease outbreaks. The BaRD is a database that catalogs and classifies epidemiological model-specific information. The SWAP is an app to provide a context for a rapidly unfolding event through graphical visualization.



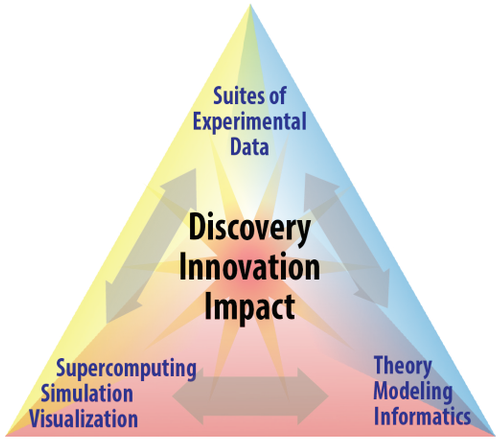
PARAVIEW is an open source visualization tool developed for the weapons program that has successfully been used to visualize output from large-scale simulations, including Episims.



Next Steps

Over the past two decades the scientific method has evolved from a close integration of experimental and theoretical sciences to include the computational and information sciences. In large measure this evolution has been driven by complexity of the science enterprise and by the sheer volume, velocity, and variety of data we now have the tools to acquire.

Making sense of this expanded scope in data requires sophisticated integration of computational sciences, simulation, and visualization with theory, modeling and experiment. This powerful approach has been applied with great success in areas such as cosmology, nuclear physics, and climate change. Our intent is to transition lessons learned in these other areas to biosurveillance.



On a concrete level, our next steps as an organization will be to refine our internal strategy, build partnerships, and invest in the people, equipment, and capabilities required to make our plan a reality. The goals articulated in the previous sections will provide the framework for this process.

The Laboratory welcomes participation from internal groups, teams, and individuals and we are constantly seeking to build new partnerships with external organizations. Should you wish details on the various aspects of the plan mentioned here, please contact us.

Planning is an evolutionary process, which means that a detailed document is often obsolete as soon as it is printed. However, the broad strokes of a plan should not change frequently, and this overview of the Laboratory's intent for biological signatures research is intended to be general enough that it will remain serviceable for several years.

The Science of Signatures is overseen by the Chemistry, Life, and Earth Science Directorate Office.

Chemistry, Life, and Earth Sciences Directorate Office
Phone: (505) 606-2266
Mail Stop: F629
Los Alamos National Laboratory, 87545

The biological signatures component of the Science of Signatures is overseen by Bioscience Division and Theoretical Division in conjunction with the other divisions and program entities listed on page 7.

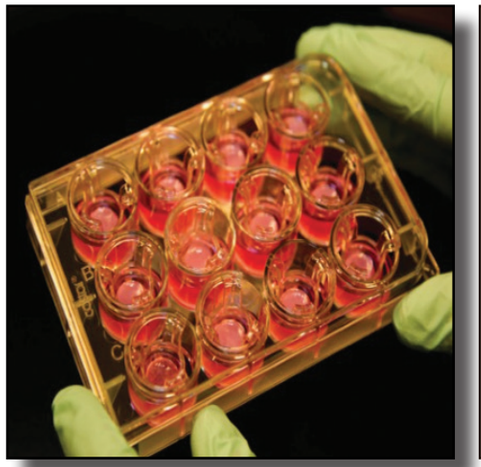
Bioscience Division
Phone: (505) 667-2690
Mail Stop: M888
Los Alamos National Laboratory, 87545

Theoretical Division
Phone: (505) 665-4055
Mail Stop: B210
Los Alamos National Laboratory, 87545

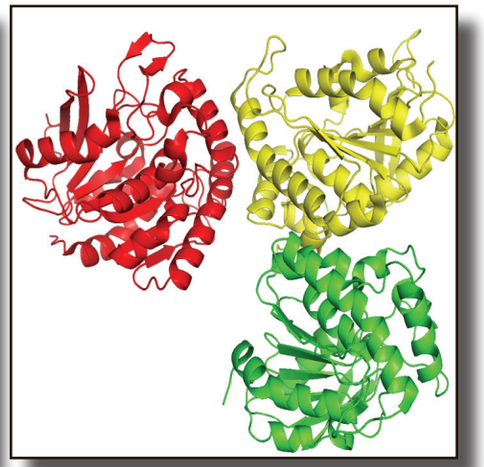
The full Science of Signatures strategic plan, of which this plan is a subset, is available from any of these offices.

This document was generated by a subset of the Biological Signatures planning team. Primary authors are Basil Swanson (B-10), Benjamin McMahon (T-6), Nicholas Hengartner (T-6), Goutam Gupta (B-10), Harshini Mukundan (C-PCS), John Dunbar (B-11), Alina Deshpande (DSA-3).

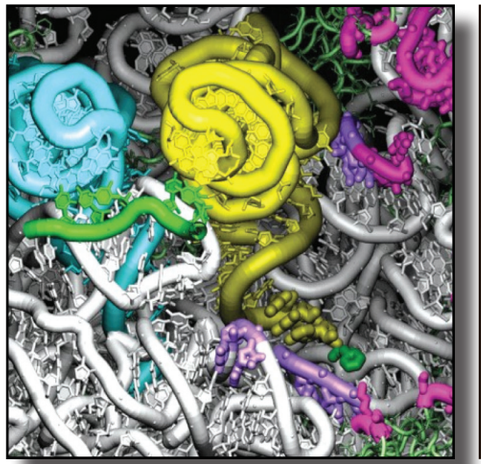
Document preparation by Josh Smith and Rebecca McDonald, ADCLES communications team.



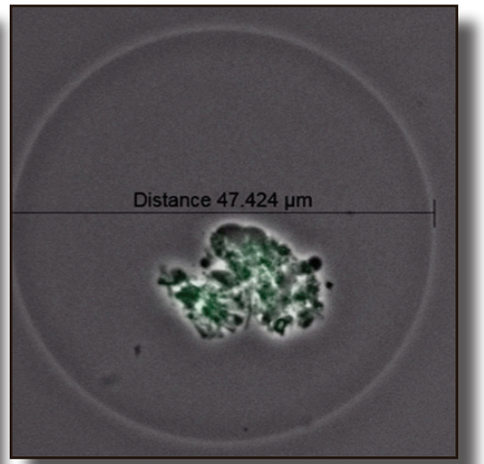
Testing platforms: engineered human tissues are being designed to replace human and animal tissues.



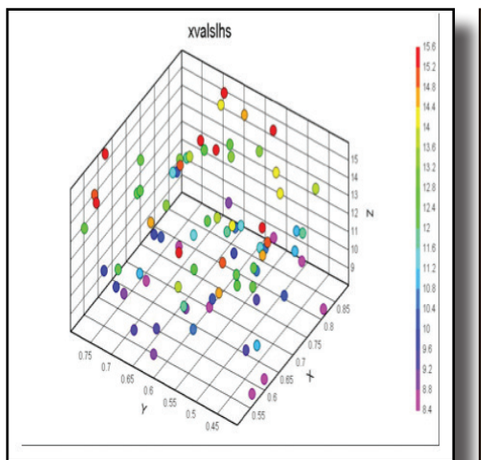
Novel detection: Ag85 is a three protein complex that is showing promise as an early indicator of tuberculosis infection.



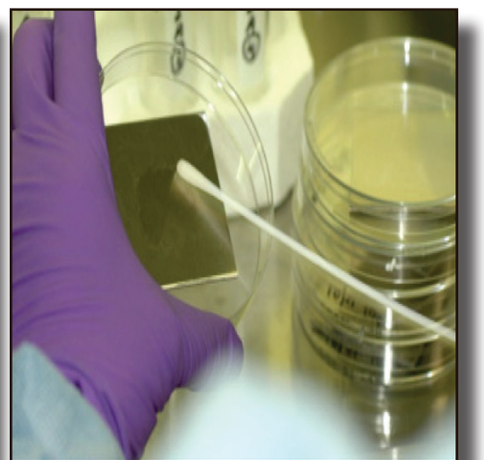
Modeling: Ribosome simulations to look for new antibiotics.



Tools: culturing a complete genome from single cells.



Uncertainty quantification for flu: an orthogonal array based latin-hypercube sample allows more accurate UQ.



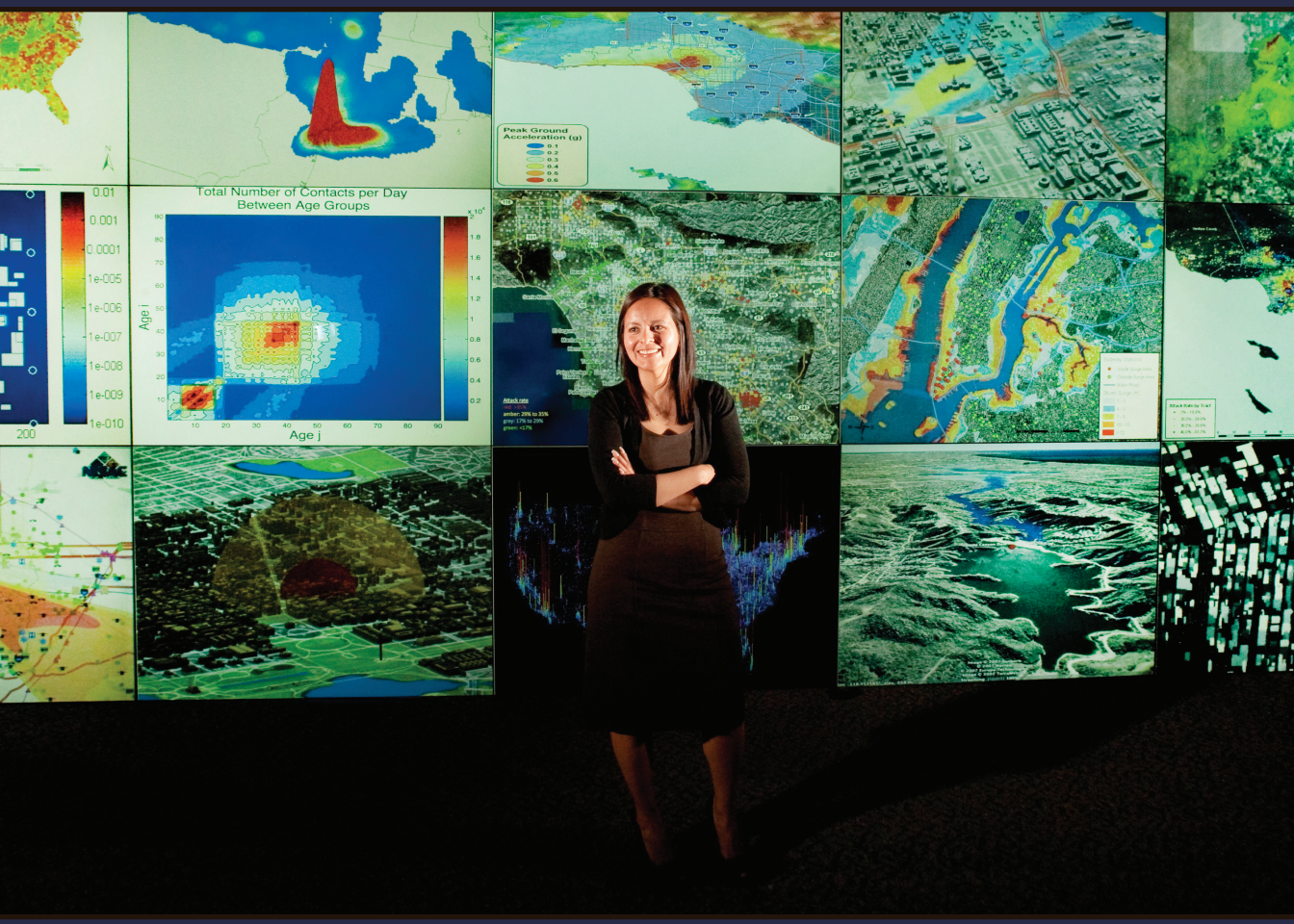
Novel detection: Various new methods of pathogen analysis are under development.

First printing, January 2014.

Los Alamos National Laboratory, an affirmative action/equal Opportunity employer, is operated by Los Alamos National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396.

A U.S. Department of Energy Laboratory

LA-UR-14-20499



Sara Del Valle, an epidemiologist with Defense Systems and Analysis Division, stands in front of the PowerWall, which is a stereoscopic theater featuring 24 synchronized projectors and screens that project high resolution, three-dimensional models. The PowerWall is used for a wide variety of modeling and simulation projects.

Los Alamos NATIONAL LABORATORY

National Nuclear Security Administration