IQT
QUARTERLY
VOL. 7 NO. 3 WINTER 2016

Mission Possible:
Quenching Epidemics

IQT
IN-Q-TEL

IQT Quarterly is a publication of In-Q-Tel, Inc., the strategic investment organization that serves as a bridge between the U.S. Intelligence Community and venture-backed startup firms on the leading edge of technological innovation. *IQT Quarterly* advances the situational awareness component of the IQT mission, serving as a platform to debut, discuss, and debate issues of innovation in the areas of overlap between commercial potential and U.S. Intelligence Community needs. For comments or questions regarding IQT or this document, please visit www.iqt.org, write to iqtquarterly@iqt.org, or call 703-248-3000. The views expressed are those of the authors in their personal capacities and do not necessarily reflect the opinion of IQT, their employers, or the Government.

EDITORIAL

IQT Quarterly, published by In-Q-Tel, Inc.
Editor-in-Chief: Adam Dove
Theme Editor: Kevin P. O'Connell
Contributing Editor: Carrie Sessine

Design by Lomangino Studio LLC
Printed in the United States of America

IQT
QUARTERLY

Identify. Adapt. Deliver.

TABLE OF CONTENTS

ON OUR RADAR

IQT IN·Q·TEL

Defeating Infectious Disease

By Kevin P. O'Connell

Infectious disease has been a topic of nearly constant media attention in the last two years due to the outbreak of Ebola virus in West Africa. The outbreak, which has continued to smolder long into 2015 (and likely through 2016), has raised questions that are central to our broader concerns about infectious disease both overseas and in the United States (and illustrate further that the distinction between *over there* and *here* are largely illusory). Why do outbreaks of infectious disease occur? Can we predict them? How do they spread? How can we respond to outbreaks more effectively? What is the role of technology in this response? In this article we consider how far we have come in understanding infectious disease, point out current issues, and identify technology trends that will drive the next generation of solutions.

	YEAR		
Rank	**1850**	**1900**	**2010**
1	*Tuberculosis*	*Influenza and pneumonia*	Heart disease
2	*Dysentery/ diarrhea*	*Tuberculosis*	Cancer
3	*Cholera*	*Diarrhea*	Chronic lung disease
4	*Malaria*	*Heart disease*	Stroke
5	*Typhoid Fever*	*Stroke*	Accidents
6	*Pneumonia*	*Kidney disease*	Alzheimer's disease
7	*Diphtheria*	*Accidents*	Diabetes
8	*Scarlet Fever*	*Cancer*	Kidney disease
9	*Meningitis*	*Senility*	*Pneumonia/ Influenza*
10	*Whooping Cough*	*Diphtheria*	Suicide

Table 1 | Leading causes of death in the U.S. in 1850, 1900, and 2010. Causes due to infectious disease are given in italics. (Sources: U.S. Centers for Disease Control and Prevention, *Journal of the American Medical Association*, and the Reuben Fleet Science Museum.)

We've Accomplished Much...

Understanding the root cause of a problem is typically the beginning of finding a lasting solution. While we have understood since antiquity that some diseases are contagious, our grasp of the germ theory of disease began only about 150 years ago. In the late 1800s, pioneers like Louis Pasteur, Ignaz Semmelweis, Robert Koch, and Paul Ehrlich began to uncover the links between microorganisms and disease. Their discoveries laid the foundations for modern vaccines, established the essential role for public sanitation, and proved the feasibility of effective antimicrobial drugs. Almost immediately, their efforts began saving lives while dispelling prior misconceptions of the cause of illnesses (e.g., bad air, imbalanced humors, and spontaneous generation). The impact of their early work is apparent when comparing the leading causes of death in the U.S. in 1850 and in 1900 (Table 1).

In 1850, all 10 leading causes of death in the U.S. were infectious diseases. By 1900, death rates overall had decreased, especially among children. These improvements were largely due to improvements in

public sanitation reducing the spread of disease, and improved nutrition which strengthened resistance to infection once it happened. Between 1900 and 2010, further improvements in nutrition, sanitation, water quality, the advent of antibiotics, and widespread childhood vaccination further increased life expectancy and removed nearly all infectious disease from the leading causes of mortality (Figure 1).

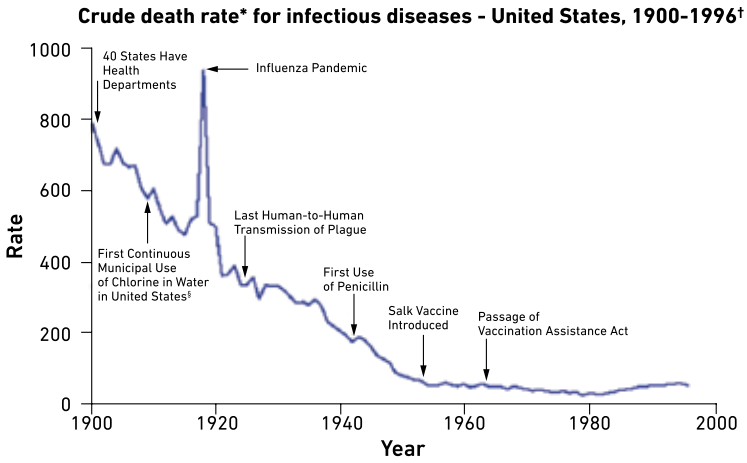
...But Much Is Left to Do

So, why do we remain concerned about infectious disease? Three clues to answering this question can also be found in the data in Table 1 and Figure 1.

First, nearly all of the diseases listed are caused by bacteria. Note that in 2010, only influenza (flu), which is caused by a virus, remains a top cause of death in the U.S. Antibiotics are effective against bacteria (but not viruses) and effectively drove down the incidence of tuberculosis and scarlet fever. They are also effective against cases of typhoid, whooping cough, diphtheria, and cholera where vaccination has not prevented illness (it is worth noting that vaccination against these illnesses is so widely practiced and effective that most modern physicians have never seen a case). Conversely, many of the diseases of greatest concern today are caused by viruses, for which there are few available therapeutics. Some viruses (notably HIV) have proven particularly difficult to prevent with a vaccine.

Influenza remains a concern globally for several additional reasons. It has a vast animal reservoir (large wild bird populations) which cannot be cured of the virus by any practical means. In recent years, highly pathogenic strains (such as H5N1) have been observed in birds; infrequent H5N1 infections in humans raise the concern that further mutations of this strain could make H5N1 more easily transmitted among humans, resulting in a strain as virulent as the 1918 strain (Figure 1). While flu in humans can be prevented by vaccination, strains of flu virus change annually (due to mutation of a gene that encodes a virus surface protein), and predictions of the vaccine strain in advance of flu season each year vary in their accuracy. Much work is now focused on developing vaccines based on parts of the flu virus that do not change, and some approaches to this problem show promise.

Second, the data are specific to the U.S. Globally, tuberculosis and malaria remain public scourges for reasons that range from economic (the cost of



* Per 100,000 population per year.
† Adapted from Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. JAMA 1999:281;61-6.
§ American Water Works Association. Water chlorination principles and practices: AWWA manual M20. Denver, Colorado: American Water Works Association, 1973.

Figure 1 | Decrease in overall death rate in the U.S. since 1900.

antibiotics), and environmental (malaria is endemic mostly in the tropics; in the U.S. it was eliminated through controlling mosquitos) and through public sanitation (diarrheal diseases are primarily transmitted by contaminated water supplies).

Third, the Table 1 data begin in 1850, a time when the world population was one-sixth of today's population (1.2 billion in 1850 versus 7.3 billion now), and our larger population moves more often, and much faster, than people did in 1850. In 1850, an ocean crossing took days to weeks, whereas people routinely travel the globe today in a matter of hours. As the human population grows, more people are competing for space with animal populations in areas that were previously considered remote. These animals (primates, bats, birds, rodents, and others) are the reservoirs for pathogens both known (influenza and Ebola) and those that, until recently, we had not seen (viruses that cause MERS and SARS are recent examples).

In addition to these points, we have begun to find the limitations of even our most powerful tools. Antibiotics can lose their efficacy as bacteria gain the ability to resist them. Some members of the public have resisted the use of vaccines, reducing the herd immunity that protects the wider population and allowing the resurgence of some diseases. So, for all these reasons, infectious diseases, especially those caused by viruses, continue to cause concern.

The Role of Technology

In the light of challenges like the 2014-2015 outbreak of Ebola, the ongoing MERS virus outbreak, the continuing threat of seasonal and highly pathogenic flu, and the problem of multiple drug-resistant bacteria, it's tempting to adopt a gloomy outlook. However, we stand now at a time when our knowledge and capabilities in biology and engineering offer tools that, with sufficient resources and will, can greatly improve our ability to combat infectious disease. These tools, including genome sequencing and genome-scale DNA synthesis (whose costs are falling faster than Moore's Law), synthetic biology, and big data analytic methods, plus our exploding understanding of the immune system and ability to engineer complex devices at low cost, are enabling advances in:

Diagnostics: Advances in microfluidics, material science, low-power electronics, and the Internet of Things are making it possible to drive down the cost of fast, sensitive, and specific diagnostic devices. Challenges in the business model for diagnostics remain, however.

Vaccines: Synthetic biology, better design tools, and nanomaterials are enabling vaccines against diseases that have resisted traditional methods.

Antiviral Medications: Our experience with HIV, among other diseases, shows that effective antiviral drugs can be developed; the challenge now is to learn to accelerate the process.

Antibiotics: The massive sequencing of bacterial genomes has helped identify new drug targets.

Beyond biology, we are also learning to use the growing power of mobile devices to connect public health workers, logisticians, analysts, scientists, and decision makers. Big data methods are being developed that will help convert raw field information into action — because now more than ever, data can move faster than people. The promise of this convergence of biotechnology and IT is that we may be able, hopefully sooner than later, to move diagnostics, vaccines, treatments, and data faster than the spread of infectious disease.

The national security implications of a rapidly spreading infectious disease, whether naturally occurring or deliberately released, are evident, and inherently relevant to the missions of IQT's customers. Outbreaks are transnational by nature, and have high potential to be politically, socially, and economically destabilizing. This issue of the *IQT Quarterly* is focused on identifying potential solutions to technology gaps in outbreak response, and introduces BiologyNext, the IQT Lab devoted to improving our ability to detect, identify, and quench such outbreaks. Q

Kevin P. O'Connell, Ph.D., *is Vice President with In-Q-Tel's Field Deployable Technologies Practice. He has been with IQT since 2007. His 30-year career began in applications of molecular microbiology to problems in agriculture, and progressed to R&D projects in biological defense. O'Connell was a scientist and principal investigator with a Department of Defense laboratory for 10 years, where he focused on pathogen detection and genetic characterization. He is the author of more than 50 peer-reviewed journal articles, book chapters, and other publications, edited a book on biological defense, and holds eight patents. O'Connell holds a B.S. in Life Sciences from MIT and an M.S. and Ph.D. in Bacteriology from the University of Wisconsin-Madison.*

A Look Inside

This issue of the *IQT Quarterly* examines converging technology areas with the potential to quench infectious disease outbreaks and their impact on the Intelligence Community's biodefense landscape.

Tara O'Toole and Stephanie Rogers of IQT Lab BiologyNext introduce the pressing national security problem of infectious disease epidemics and the concurrent biorevolution — the convergence of computational and analytical technologies with decades of progress in biology. BiologyNext aims to better understand infectious disease epidemics and construct an architecture of technologies that could be used to rapidly detect and quench them.

Peter Daszak of the EcoHealth Alliance discusses common trends for predicting pandemics and challenges to preventing them. The rising frequency of infectious diseases, their increasing geographic spread, and expanding impact should make overcoming pandemics an international priority.

Next, the *IQT Quarterly* interviews Larry Madoff, Editor of the Program for Monitoring Emerging Diseases (ProMED-mail). ProMED-mail reports on outbreaks and disease emergence, providing early warning information to a global audience and facilitating informed discussions in real time. Madoff discusses the service's reporting processes, the broader community of infectious disease reporting, and the future of outbreak detection and response.

Lenny Moise, Sarah Beseme, and Anne De Groot of EpiVax discuss recent computational advances that accelerate vaccine design in response to biological threats. These tools make it possible to design vaccines in the shortest time possible once the DNA sequence from the emerging infectious disease is available, and will be critical for nations to rapidly respond to outbreaks of known and novel pathogens.

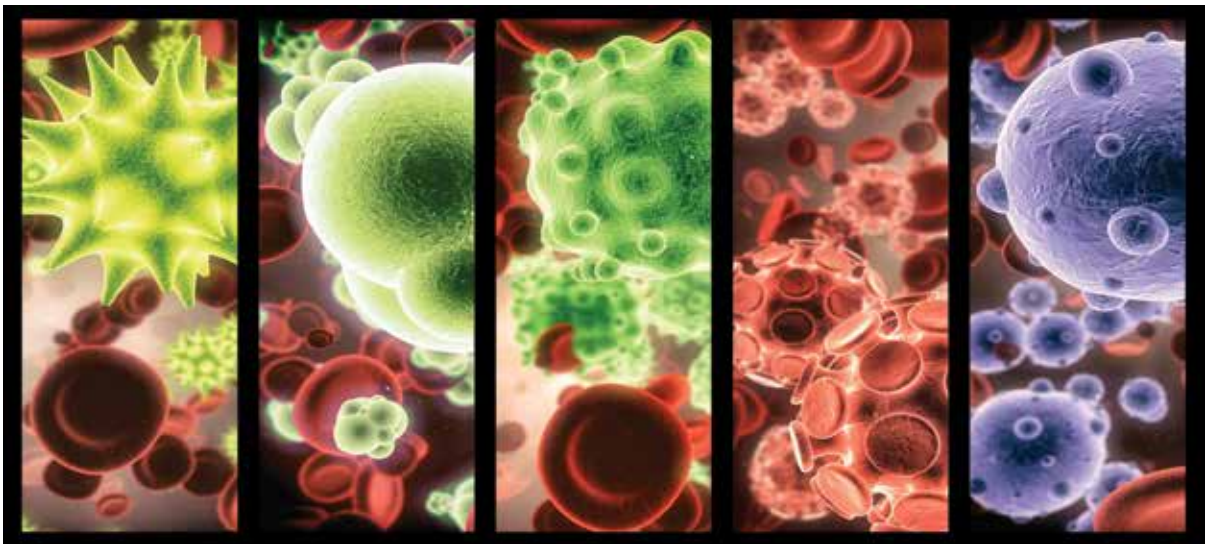
Mark Lim of The Gates Foundation provides insights into diagnostic sample challenges. He describes an engineering workflow from human sample to molecular answers that can mitigate data variability and bias, and provide value in the quenching of disease outbreaks.

Finally, we close the issue with a technology overview from IQT portfolio company Quanterix, whose digital approach to immunoassays allows single molecules to be counted and enables a massive increase in sensitivity. Combined with full automation, this sensitivity provides unprecedented insight into disease detection, diagnosis, and patient treatment.

Beyond the topics covered in this issue, there is a continued global conversation around infectious disease outbreaks and the converging technologies that can be harnessed to quench their destabilizing impact. Identifying relevant technology gaps and solutions will be critical for national security. Q

IQT Lab BiologyNext and the National Security Implications of 21st Century Life Sciences

By Tara O'Toole and Stephanie Rogers



BiologyNext, an IQT Lab, will explore a complex and increasingly urgent problem: how can we rapidly detect and quench epidemics of infectious disease — whether they arise from natural causes or acts of bioterror?

Since 2013, IQT has launched four strategic initiatives, known collectively as IQT Labs. These initiatives are intended to address complex, urgent national security problems as they intersect with disruptive, game-changing evolutions in science and technology. IQT Labs will pursue collaborations among government customers, innovative private sector partners, and academia in an effort to understand, illustrate, and demonstrate emerging technologies and their potential roles in national security.

The persistent — and increasing — risk of infectious disease epidemics and the potential for catastrophic bioterror attacks are evident national security concerns. Meanwhile, extraordinary advances are being made in the life sciences and biotechnologies. The capabilities produced by the digital revolution — e.g., sensors, advanced analytics, and mobile communications — are converging with our ability to comprehend and manipulate the parts, circuits, and operating systems of living organisms and biological systems. BiologyNext will seek to exploit the expanding understanding of how epidemics arise and unfold, and the emerging array of powerful biotechnologies to construct a technology architecture for biodefense. This architecture of extant and emerging capabilities will serve to map how, given the appropriate will, imagination, and resources, we might significantly improve epidemic detection and response.

The Stubborn Challenge of Disease Outbreaks

Throughout history, infectious disease outbreaks have demonstrated their potential to cause vast suffering and societal disruption. Smallpox, which was eradicated in 1980 after a long campaign led by the World Health Organization, killed 300-500 million people in the 20th century alone — more deaths than resulted from all the wars of that bloody period.[1] The 1918 influenza pandemic killed 60-100 million people worldwide, including 675,000 Americans.[2] In comparison, the estimated combined death toll from the Hiroshima and Nagasaki bombings is 185,000.

While vaccines and public hygiene have greatly reduced the global death toll from infectious disease, the problem of disease outbreaks persists. New strains of influenza emerge annually, requiring public health

authorities to predict in advance what strains to include in each year's vaccine production. Newly emergent diseases such as SARS (2003), MERS-CoA (2012), and Ebola (several outbreaks since the 1970s) demonstrate that the world harbors a wide variety of pathogenic viruses with many yet undiscovered. The outbreak of Ebola, which began in West Africa in December 2013, demonstrates the extraordinary human, economic, and social impact of an emergent epidemic disease. This outbreak has thus far infected more than 28,000 West Africans, resulting in more than 11,000 deaths (as of November 2015).[3] It is estimated that just as many deaths, if not more, have resulted from malaria left untreated during the peak of the epidemic.[4,5] The economic damage, including both the costs of response (the U.S. Government has provided over $2.3 billion as of November 2015) and the direct and indirect impacts on involved countries are in the tens of billions of dollars.[3] Social costs include the near total destruction of health systems in Liberia, Guinea, and Sierra Leone; the loss of hundreds of heroic health professionals; and continued stigmatization of Ebola survivors in West Africa and of African immigrants around the world.[6,7]

For many reasons, our world is entering a period in which large-scale, lethal epidemics are likely to become more frequent, affect more people, and spread faster and farther than has been the case historically.[8] For example:

Climate: The warming climate is changing the geographical distribution of animals and disease vectors, introducing microbial threats from tropical and sub-tropical areas into temperate zones.[9] Melting permafrost has exposed previously unknown viruses.

Population: Global urbanization has produced dozens of megacities with upwards of 10 million inhabitants each, most of whom live in conditions of poor nutrition, inadequate sanitation, and limited access to health care, creating optimal conditions for the incubation and spread of pathogens. The growing human population and ensuing economic pressures are thrusting people into once remote ecosystems where they are in contact with previously unknown microbes.[10]

Agriculture and Land Use Changes: The increasingly industrialized production of meat and poultry gathers vast numbers of animals together in close contact with humans and with each other, producing potential for animal and human infections. Moreover, the overuse of antibiotics in agriculture is a major source of the increase in antibiotic resistance, which results in 23,000 U.S. deaths annually.[11]

Mobility: Modern trade and travel patterns ensure the continuous movement of people, animals, plants, and microbes at the speed of jet airliners. Political disruption is causing the migration of people in numbers without historical precedent. According to the United Nations, more than 43 million people are currently displaced due to use of force or persecution, half of them children.[12]

The Dual Use Dilemma of Biotechnology

The threat of disease outbreaks is not limited to naturally occurring pathogens. The use of disease in warfare dates to antiquity, and has been studied intensively by nation-states and non-state actors since the early years of the 20th century. Since then, advances in biotechnology and pharmacology have steadily lowered the bar to engineering, producing, and disseminating pathogenic bacteria and viruses. The biological techniques that are essential to creating new medicines and other valuable products are also useful to state and non-state malefactors, and most of these pharmaceutical innovations have been published.[13] As is the case with most technologies, biotechnologies become cheaper, more accessible, and easier to use as they mature. Moreover, because the knowledge and materials needed to build a biological weapon are dual-use in nature and essential for legitimate bioscience work, it is extremely difficult to detect efforts to develop bioweapons.

Shifting the Advantage to Biodefense

BiologyNext will work with others to design a technology architecture for effective biodefense, to significantly improve the ability to rapidly detect and quench destabilizing epidemics, whether natural or engineered. To do so, we will draw upon technological capabilities emerging from the life sciences, data science, communications, and information technology.

Raising Awareness of the Biorevolution and the Feasibility of a Robust Biodefense

BiologyNext will pursue several approaches to develop and communicate the nature of the epidemic threat; to identify key points at which emerging technologies could be used to detect and interrupt the spread of an epidemic and lessen its impact; and to describe the existing and emerging technological capabilities to accomplish this. BiologyNext is exploring a variety of platforms which could vividly convey the nature of the epidemic threat and the potential offered by biological and other technologies to shift the advantage to biological defense.

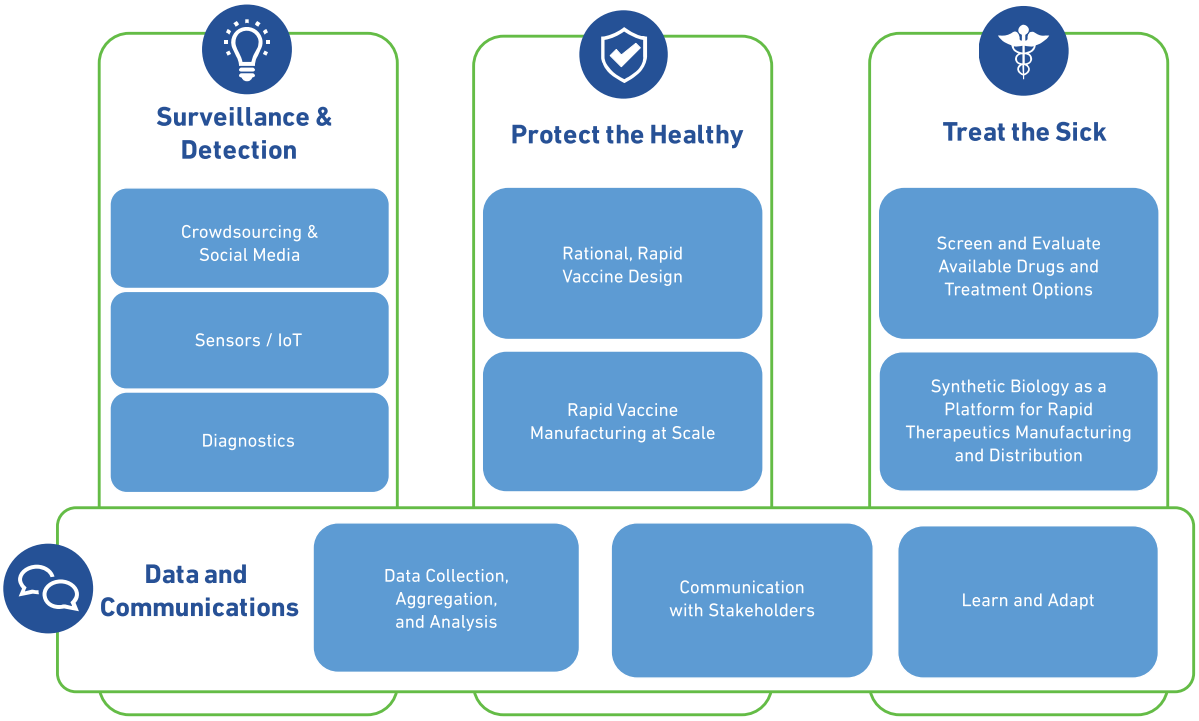


Figure 1 | *BiologyNext* technology architecture, showing components required to identify, characterize, and quench infectious disease outbreaks.

BiologyNext is also convening a series of roundtables consisting of subject matter experts from industry, academia, and government to discuss the technology developments that are most vital to detect and interrupt disease outbreaks. These roundtable discussions are intended to convey specific information about current capabilities, provoke strategic thinking about biodefense and emerging technologies, and build a community of government personnel and technical experts from industry and academia who can continue the conversation about biodefense beyond BiologyNext. The summaries from these discussions will be captured and made available in the form of white papers or similar publications, and will be used to inform the development of a biodefense technology architecture.

Building a Technology Architecture for Effective Epidemic Detection and Response

The BiologyNext technology architecture is an evolving framework of technologies which, collectively, could substantially improve epidemic detection and response (Figure 1). Examples of such technologies include approaches to rapid detection and diagnosis of previously unknown "Agent X" pathogens; rational, rapid vaccine design and manufacture at scale; and effective means of obtaining situational awareness, communication, and real-time learning during epidemics. We are hopeful that the process of mapping key functional elements of epidemic detection and management, and assembling a specific architecture of technologies to meet these needs, will provide a set of avenues focused on ending destabilizing epidemics.

IQT Lab Challenges: Selected Proofs of Principle

BiologyNext plans to pursue several "challenge" projects designed to demonstrate proof-of-principle for key technologies or methods within the overall architecture. For example, situational awareness during disease outbreaks typically lags reality by days or weeks. Decades of work and billions of dollars have been spent attempting to build useful surveillance systems. A potential challenge is the construction of a test bed in which potential approaches could be evaluated for efficacy, cost, and practicality. Another challenge might be the exploration of machine learning and other big data analytic techniques to develop predictive algorithms that could be used to rapidly identify the epitopes essential to a viable vaccine against a new and unknown pathogen. It is also expected that BiologyNext products

will inform and lead future IQT investments in the life sciences, particularly where such investments embody synergies with IQT customer requirements.

The Digital Revolution Changed the World; The Biorevolution Will Transform It

No matter their origin, infectious diseases pose significant, potentially existential, threats to the health, well-being, and economic competitiveness of our nation. While efforts have been made over the last few decades to build a national response capability, recent experience has demonstrated that we are substantially unprepared to handle significant outbreaks, especially of previously undescribed pathogens. The BiologyNext architecture will show how the tools and capabilities emerging from the convergence of the digital revolution and the rapidly evolving biorevolution offer an unprecedented opportunity to substantially improve our ability to detect and quench lethal, large-scale epidemics. Q

***Tara O'Toole, M.D., M.P.H.**, is an Executive Vice President at In-Q-Tel and Director of IQT Lab BiologyNext. Prior to IQT, O'Toole was the Under Secretary for Science and Technology at the Department of Homeland Security. In this role, she was responsible for scientific and technological research designed to provide new security and resilience innovations. Previously, she was the CEO and Director of the Center for Biosecurity at the University of Pittsburgh Medical Center (UPMC), an independent organization dedicated to improving the country's resilience to major biological threats. O'Toole is internationally known for her work on biosecurity and on health and safety issues related to the U.S. nuclear weapons complex. O'Toole has a bachelor's degree from Vassar College, an M.D. from George Washington University, and a Master of Public Health degree from Johns Hopkins University.*

***Stephanie Rogers, Ph.D.**, is Deputy Director of IQT Lab BiologyNext. Prior to this role, she was a Program Manager at In-Q-Tel, where she managed biological and chemical investments in IQT's field deployable technologies portfolio. She was previously invited by the FBI to serve as a fellow in the Oak Ridge Institute for Science and Education (ORISE) Visiting Scientist Program. In this role, she managed and conducted research projects in human forensics and microbiology for the FBI's Counterterrorism and Forensic Science Research Unit. Rogers holds a Ph.D. in Plant Pathology and a B.S. in Biochemistry and Molecular Biology from Oklahoma State University.*

REFERENCES

1 F. Fenner, D.A. Henderson, I. Arita, Z. Ježek, I.D. Ladnyi. Smallpox and its Eradication. WHO. Geneva. 1988

2 Johnson NP, Mueller J. Updating the accounts: Global mortality of the 1918–1920 "Spanish" influenza pandemic. Bull Hist Med. 2002. 76:105–115

3 USAID West Africa Ebola Outbreak Fact Sheet #3-11-06-2015(https://www.usaid.gov/sites/default/files/documents/1866/west_africa_fs03_11-06-2015.pdf)

4 Walker PGT, White MT, Griffin JT, Reynolds A, Ferguson NM, Ghani AC. Malaria morbidity and mortality in Ebola-affected countries caused by decreased health-care capacity, and the potential effect of mitigation strategies: a modelling analysis. Lancet Infect Dis 2015. 15:825-32

5 Plucinski MM, Guilavogui T, Sidikiba S, Diakité N, Diakité S, Dioubaté M, Bah I, Hennessee I, Butts JK, Halsey ES, McElroy PD, Kachur SP, Aboulhab J, James R, Keita M. Effect of the Ebola-virus-disease epidemic on malaria case management in Guinea, 2014: a cross-sectional survey of health facilities. Lancet Infect Dis 2015. 15:1017-1023

6 Mullan Z. The Cost of Ebola. Lancet Glob Health 2015. 3:e423

7 World Bank. Update on the Economic Impact of the 2014 Ebola Epidemic on Liberia, Sierra Leone, and Guinea. December 2, 2014. http://www.worldbank.org/content/dam/Worldbank/document/Economic%20Impact%20Ebola%20Update%202%20Dec%202014.pdf

8 Morse SS. Factors in the Emergence of Infectious Diseases. Emerging Infectious Diseases. 1995. 1:7-15

9 Altizer S, Ostfeld RS, Johnson PTJ, Kutz S, Harvell CD. Climate Change and Infectious Diseases: From Evidence to a Predictive Framework. Science 2013. 341:514-19

10 Neiderud, CJ. How urbanization affects the epidemiology of emerging infectious diseases. Infect Ecol Epidemiol 2015. 5:27060

11 Liverani M, Waage J, Barnett T, Pfeiffer DU, Rushton J, Rudge JW, Loevinsohn ME, Scoones I, Smith RD, Cooper BS, White LJ, Goh S, Horby P, Wren B, Gundogdu O, Woods A, Coker RJ. Understanding and Managing Zoonotic Risk in the New Livestock Industries. Environ Health Perspect 2013. 121:873-77

12 United Nations High Commissioner for Refugees. Global Trends: Forced Displacement in 2014. http://unhcr.org/556725e69.pdf

13 Volkin DB, Hershenson S, Ho RJY, Uchiyama S, Winter G, Carpenter JH. Two Decades of Publishing Excellence in Pharmaceutical Biotechnology. J Pharm Sci 2015. 104:290-300

14 Atlas RM, Dando M. The Dual-Use Dilemma for the Life Sciences: Perspectives, Conundrums, and Global Solutions. Biosecur Bioterror 2006. 4:276-286

Five Major Challenges for Pandemic Prediction and Prevention

By Peter Daszak

Dissecting the Anatomy of a Pandemic

Pandemics (diseases that spread globally) are rare events that are often devastating, causing substantial mortality and economic damage. Just like hurricanes or earthquakes, efforts to understand the origins of pandemics and predict their emergence would help reduce their impact and ultimately prevent them. However, unlike earthquakes or hurricanes, our efforts to understand the causes, patterns, and origins of pandemics are only just beginning. Here I highlight recent advances in disease ecology, virology, and biogeography that move us towards these goals. I also identify five critical questions that, if answered, will greatly enhance our ability to predict and prevent pandemics.

Predicting pandemics first requires analyzing trends and common themes in their emergence

Over the past few decades we have learned a great deal about the anatomy of a pandemic. Most pandemics originate as zoonoses (diseases from animals, mainly wildlife). In fact, every one of the true pandemics of the last 50 years has either originated entirely within a wild animal species (e.g., SARS originating in bats) or contains genes derived from wild animal viruses (e.g., pandemic influenza A H1N1/09 virus).[1] Because most pandemics are caused by viruses, this article will focus on them. Pandemics emerge through a complex interplay among socioeconomic, ecological, and biological factors. This process contains at least three distinct stages.[2]

- First, potentially pandemic pathogens exist in their natural wildlife reservoir. Changes to land use or other environmental changes bring people into increased contact with wild animal hosts, or perturb natural host-pathogen dynamics to increase the risk of viral transmission from wildlife to people (i.e., "spillover").[3]
- Second, spillover to human occurs repeatedly, either directly from a wildlife host or via domesticated animals. Some spillover events cause small chains of human-to-human transmission.
- Third, the virus achieves sustained human-to-human transmission, expands its geographical range, and moves internationally via travel and trade networks. This stage is pandemic emergence: international spread with sustained transmission across large swathes of the planet.

Each stage of emergence is driven by different socioeconomic, ecological, and biological factors that push pathogen dynamics through emergence,

amplification, and spread. However, these processes are poorly understood because they are complex; elucidating their mechanisms requires collaboration across many disciplines (e.g., demographers, virologists, and wildlife biologists). However, general trends can be identified.

Emerging infectious diseases are increasing in frequency, pandemic potential, and impact

The literature on emerging infectious diseases (EIDs) is growing. Does this growth reflect an actual increased threat of pandemics, or simply better reporting of outbreaks? To test this, we developed a database of EID events (expanded from a published list of EIDs) which we define as the first temporal emergence of a pathogen in a human population or the point at which a previously known disease became classified as emerging due to increased incidence or other factors.[4,5] We collected and analyzed data on the location and time at which all EIDs since 1960 emerged, and a series of associated ecological, biological, and socio-demographic drivers of disease emergence. Our spatial and temporal regression analyses showed that the frequency of EID events has increased over time, peaking between 1980 and 1990. This peak was associated with increased susceptibility to infection due to the HIV/AIDS pandemic. Like Taylor et al. (2001), we found that zoonoses comprised the majority of EID events (60.3 percent), and that almost 71.8 percent of zoonotic EIDs were from wildlife (43.3 percent of all EID events).[4] Furthermore, zoonoses from wildlife were increasing as a proportion of all EID events — between 1990 and 2000, 52 percent of EID events were zoonoses with a known wildlife origin. We attempted to correct for increasing infectious disease reporting by including in our regression model the number of articles published in the *Journal of Infectious Diseases* (which gives a crude measure of research effort for infectious diseases generally, not just EIDs). Controlling for the frequency of reporting further supported our conclusions that EID events are becoming more common, that zoonoses comprise the majority of EID events, and that zoonoses are rising significantly faster as a proportion of all EID events.

Identifying hotspots for pandemic disease emergence

Using geographic data in our EID database, we tested associations between subsets of EID events

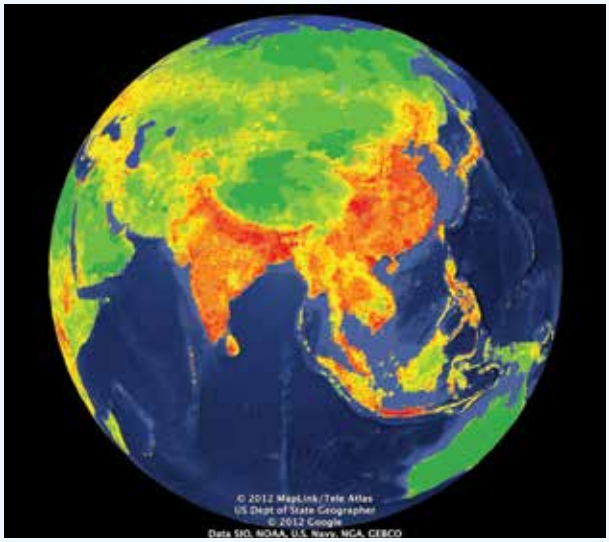


Figure 1 | Heat map of risk of a zoonotic disease of wildlife origin emerging in people; warmer colors reflect increased risk of EID events (EID hotspots). Because pandemics are mainly zoonotic in origin, this map acts as a potential basis for future targeted surveillance and the pre-empting of potential pandemics.

(drug-resistant and vector-borne pathogens, and zoonoses with wildlife and non-wildlife origins) and a few hypothesized drivers. While each category of EID event was associated with different outbreak drivers and geographic patterns of origin, all were strongly associated with human population density. This finding suggests that the presence and number of people and the changes they make to a landscape are key risk factors for emerging diseases. We also found that land use (e.g., urbanization, change in land cover) correlates significantly with EID event distribution. Zoonoses from wildlife also correlate significantly with mammal species richness; more animal species in a given location host a greater diversity of microbes. This approach provides a way to identify the relative risk of EID events (through their correlated drivers) globally. These EID hotspots tend to be lower-latitude, developing countries with high populations of people, high wildlife diversity, and lots of land use change.[5] Our hotspot maps (Figure 1) provide a first crude attempt at pandemic prediction — they identify regions likely to propagate the next EID event or Stage 1 spillover. This spatial and temporal analysis of EID events can provide a simple but powerful way to prioritize resources for global disease surveillance. The goal of this surveillance would be identifying pathogens likely

to become the next EID, or pathogens in the process of emerging. However, they also raise critical questions we need to address in our quest to prevent pandemics.

Five Critical Questions

1. How many unknown viruses are waiting to emerge?

The first step in a global program to prevent pandemics might be to survey the extent of microbial diversity in hotspot regions. Sampling wildlife and identifying all the viruses they harbor would generate a pool of potential pandemic pathogens from which to develop vaccines and other medical countermeasures. This approach is exactly the basis for a number of new programs, including the USAID Emerging Pandemic Threat (EPT) program, and research programs that target pathogen discovery in bats and other zoonotic disease reservoirs.[2, 6] However, scale is critical. If there are 30 million unknown viruses in wildlife, it will be extremely costly to identify them all, and difficult to assess which pose the greatest threats.

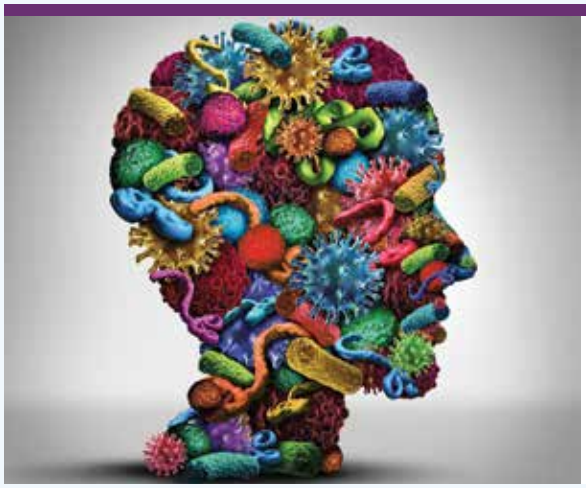
So far, only one systematic attempt has been undertaken to predict the unknown viral diversity in a single animal host species.[7] We used samples collected and tested through the USAID EPT PREDICT program, in which animals were captured, tagged, and released, and the number of recaptures of tagged individuals relative to the number of untagged individuals yielded a statistical prediction of the total number of individuals in a region. For pathogen discovery, we repeatedly sampled a large population of *Pteropus giganteus*, a bat species known to carry zoonotic viruses. From high quality samples collected from around 2,000 unique bats, we discovered 55 viruses from 9 viral families known to harbor zoonoses.[7] We estimated the total viral richness of these 9 families in *P. giganteus* to be 58 viruses (i.e., 3 not-yet discovered). Extrapolating to all 5,517 known mammal species, we estimated that there are at least 320,000 mammalian viruses awaiting discovery in these 9 viral families. Using field and lab expenses of the PREDICT program, the cost to uncover 100 percent of virodiversity in all mammalian reservoirs will be $6.8 billion, and to uncover 85 percent of virodiversity will be $1.4 billion, considering the diminishing returns of continued sampling. The latter figure is less than the cost of a single SARS-scale pandemic and, if spread over a decade, a small portion of current global pandemic prevention spending. This first attempt at estimating pathogen diversity is crude and has a number of significant assumptions. Further work on other species is critical to assess its validity.

2. In which wildlife species will the next pandemic pathogen originate?

The cost of surveying pathogens in mammalian species may be reduced by prioritizing those wild species most likely to harbor viruses that could cause a pandemic. One criterion might be close genetic relatedness to humans (e.g., non-human primates). Such animals are more likely to share with humans similar host cell receptors and viral defense mechanisms. Potentially, viruses that can infect animals related to humans are more likely to be able to infect our cells, and less likely to be quashed by our innate immunity. While the genetic, behavioral, and ecological rules that underpin these relationships are poorly understood, recent work shows that closely-related mammals are more likely to share virus species.[8] However, many mammalian genera and families have been too poorly sampled to assess their risk. Furthermore, for reasons not well-understood, this rule can break down when two unrelated species have extensive, intimate contact over long periods of time (e.g., humans and domesticated mammals). Expanding trade in wildlife may also bring more animal species from different regions into close contact with people; creating opportunities for pathogens to emerge that would normally have difficulty infecting people.

3. Can we predict the pandemic potential of a newly discovered pathogen?

Even if we increase the rate of virus discovery, how can we identify which viruses will be able to infect humans? Testing in animals or cultured human cells can predict host range and potential pathogenicity to humans, but these methods are not definitive. Other factors which may suggest disease-causing potential include the relatedness of the host species to humans, the relatedness of a virus to known human viruses, known host range, and evolutionary capacity.[2] A few heuristics can help us prioritize certain pathogens: those existing at an interface of human-wildlife contact, those closely related to known human pathogens, and whose hosts are closely related to humans are likely to pose greater risks. For close neighbors of known pathogens,



The rising frequency of emerging infectious disease events, their increasing geographic spread, and their expanding impact should make overcoming pandemic disease an international priority.

sequencing their receptor binding domains may provide a rudimentary strategy to assess potential virulence. For example, the SARS-like coronaviruses (SL-CoV) identified by our group in bats in China have varying degrees of sequence homology to the SARS coronavirus (SARS-CoV).[1, 9] We have shown that some SL-CoVs can bind directly to the human cell surface receptor for SARS-CoV, ACE2. Others have now demonstrated that chimeric viruses (SARS-CoV backbone expressing SL-CoV spike protein) can infect human cells and cause clinical signs in humanized mice.[10] Importantly, there appears to be a lack of immune cross-reactivity between SARS-CoV and our SL-CoV, suggesting that our SL-CoV has significant pandemic potential. These studies provide proof-of-concept for predictive approaches.

4. Can we predict how, and where, a new EID will spread?

The emergence of A/H1N1 influenza in 2009 highlighted how rapidly diseases can spread once they have achieved capacity for transmission. Analyses of travel and trade data have shown that modeling disease spread is relatively straightforward, and can provide accurate estimates of spread and case numbers when applied to prior outbreaks, e.g. of SARS and A/H1N1 influenza.[11, 12] This approach has been used to analyze the spread of disease vectors via shipping and likely routes of spread via airplane, and to predict the spread of ongoing events such as the MERS-CoV outbreak in Saudi Arabia.[13, 14] Models of disease spread have been used to examine the likely cause of past spreading events for A/H5N1 influenza, and predict and set policy for its likely route of introduction to the New World.[15] Finally, modeling has reduced the risk of West Nile introduction to Hawaii and the Galapagos Islands (the most likely vector is mosquitoes transported via air travel).[16]

The predictive power of such models improves with the quality of data available. For example, accurate predictions about disease spread require countries to identify and report outbreaks early once a pandemic has begun. During the 2009 H1N1 influenza pandemic, two key factors influenced the pandemic's arrival date in a given country: the country's accessibility via air travel, and the percentage of GDP per capita spent on healthcare (a proxy for testing and reporting capacity).[12] Less well-understood is the role of intra-country human movement in disease spread. New data on roads, migration, and human network connectivity will increasingly illuminate this area.

5. How do we prevent pandemics from emerging?

Even if we can identify the capacity of a novel virus for human-to-human transmission and predict how it will spread, we still lack strategies to prevent evolution of an epidemic into a pandemic. One positive development is a change in how pandemic prevention programs are funded and managed. Traditionally, outbreaks were handled by state and national agencies, which fund the World Health Organization and field laboratory networks. H5N1 influenza emerged in several small-scale outbreaks, which suggested chronic persistence

in backyard poultry farms. In response, a systems approach to pandemic prevention was developed, as well as a "One Health" collaboration of animal health, public health, and environmental agencies.[17, 18] International development agencies, which had set up programs targeting individual infectious diseases, are now actively involved in this systems approach to pandemic prevention, including support for crucial infrastructure investments, and a specific focus on collaborative One Health programs.[17] With most EID events occurring in the developing world, disease-based programs for HIV/AIDS, malaria, TB, and polio do not address the underlying flaws in public health systems that predispose locations to outbreaks of emerging infectious diseases.[19]

Future work may target the underlying drivers of disease emergence, providing economic incentives to improve practices and reduce the EID threat. For example, promoting the farming of wildlife species for consumption in place of wild-caught animals would likely reduce EID risk. This support, including better regulation and stronger agricultural institutions, should reduce the inflow of wild-caught animals and simultaneously help manage biosecurity.

Similarly, 43 percent of past EID events are attributable to changes in land use and agriculture, including logging, oil and gas, mining, and plantations. The economic impact of EIDs from land-use change is estimated to be $10-40 billion over the next 10 years, which could be considered a potential liability to extractive industries. Industrialized mining and plantation operations in EID hotspot countries are likely to be on the frontline of disease outbreaks, and are often under pressure to improve their environmental impacts. Incentives could be built into World Bank loans or concession agreements to run emerging disease impact assessments (e.g., surveillance in wildlife at a mining development site), or to mandate the building of clinics and diagnostic labs that conduct surveillance for novel EIDs at extractive sites.

Conclusion

Predicting pandemic disease emergence is difficult and complex. Overcoming pandemics will require new technological solutions, better interdisciplinary collaboration, and significant funding. However, a simple economic assessment suggests they are likely to have a substantial return on investment. With the costs of pandemics rising (between $10 and $30 billion for SARS), the relatively moderate cost (less than $10 billion) of conducting targeted surveillance, identifying novel pathogens from key wildlife species, and analyzing their potential risk becomes more attractive. There is a critical, urgent time window of around 20 years within which a global strategy to prevent pandemics needs to be implemented before the rate and cost of disease emergence expands to swamp out any possibility of control. The rising frequency of EID events, their increasing geographic spread, and their expanding impact should make overcoming pandemic disease an international priority. Q

***Peter Daszak, Ph.D.**, is President of EcoHealth Alliance, a global organization dedicated to innovative conservation science linking ecology and the health of humans and wildlife. Daszak has also consulted for other non-profit organizations and governmental agencies such as the OIE ad hoc working group on amphibian diseases, the National Academy of Sciences, U.S. Department of the Interior, International Union for the Conservation of Nature, National Institutes of Health, Australian Biosecurity CRC, DIVERSITAS, Society for Conservation Biology, and the Millennium Ecosystem Assessment. He is a recipient of the 2000 CSIRO medal for collaborative research, and his work has been the focus of extensive media coverage, including articles in The New York Times, The Wall Street Journal, The Economist, The Washington Post, U.S. News & World Report, and broadcast appearances on 60 Minutes II, CNN, ABC, NPR's Talk of the Nation, Morning Edition, and Fresh Air with Terry Gross.*

ACKNOWLEDGMENTS

This article is based, in part, on talks given to a meeting of the IOM Forum on Microbial Threats, September 24-25, 2013 and at In-Q-Tel, October 2015. Funding was provided by USAID-EPT-PREDICT, NIAID, NSF/NIH EEID, and core funding from EcoHealth Alliance.

BIBLIOGRAPHY

[1] Ge, X.-Y. *et al.* Isolation and characterization of a bat SARS-like Coronavirus that uses the ACE2 receptor. *Nature* 503, 535-538 (2013).
[2] Morse, S. S. *et al.* Prediction and prevention of the next pandemic zoonosis. *Lancet* 380, 1956-1965 (2012).
[3] Murray, K. A. & Daszak, P. Human ecology in pathogenic landscapes: two hypotheses on how land use change drives viral emergence. *Current Opinion in Virology* 3, 79-83, doi:10.1016/j.coviro.2013.01.006 (2013).
[4] Taylor, L. H., Latham, S. M. & Woolhouse, M. E. J. Risk factors for human disease emergence. *Philosophical Transactions of The Royal Society B-Biological Sciences* 356, 983-989 (2001).
[5] Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nature* 451, 990-994 (2008).
[6] Drexler, J. F. *et al.* Bats host major mammalian paramyxoviruses. *Nature Communications* 3, 796 (2012).
[7] Anthony, S. J. *et al.* A strategy to estimate unknown viral diversity in mammals. *MBio* 4, doi:10.1128/mBio.00598-13 (2013).
[8] Bogich, T. L. *et al.* in *New directions in conservation medicine* (eds A. A. Aguirre, R.S. Ostfeld, & P. Daszak) 607-618 (Oxford University Press, 2012).
[9] Li, W. *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310, 676-679 (2005).
[10] Menachery, V. D. *et al.* A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med* advance online publication, doi:10.1038/nm.3985 http://www.nature.com/nm/journal/vaop/ncurrent/abs/nm.3985.html#supplementary-information (2015).
[11] Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America* 101, 15124-15129 (2004).
[12] Hosseini, P., Sokolow, S. H., Vandegrift, K. J., Kilpatrick, A. M. & Daszak, P. Predictive power of air travel and socio-economic data for early pandemic spread. *PLoS ONE* 5, e12763 (2010).
[13] Tatem, A. J. The worldwide airline network and the dispersal of exotic species: 2007-2010. *Ecography* 34, 94-102 (2009).
[14] Khan, K. *et al.* Global Public Health Implications of a Mass Gathering in Mecca, Saudi Arabia During the Midst of an Influenza Pandemic. *J. Travel Med.* 17, 75-81, doi:10.1111/j.1708-8305.2010.00397.x (2010).
[15] Kilpatrick, A. M. *et al.* Predicting the global spread of H5N1 avian influenza. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19368-19373 (2006).
[16] Kilpatrick, A. M. Globalization, Land Use, and the Invasion of West Nile Virus. *Science* 334, 323-327, doi:10.1126/science.1201010 (2011).
[17] Bogich, T. L. *et al.* Preventing Pandemics Via International Development: A Systems Approach. *Plos Medicine* 9, doi:10.1371/journal.pmed.1001354 (2012).
[18] FAO *et al.* Contributing to One World, One Health: a strategic framework for reducing risks of infectious diseases at the animal-human-ecosystems interface. 1-67 (Food and Agriculture Organization; World Organisation for Animal Health; World Health Organization; United Nations System Influenza Coordinator; United Nations Children's Fund; World Bank, Rome, 2008).
[19] Standley, C. J. & Bogich, T. L. International development, emerging diseases, and eco-health. *EcoHealth*w10, 1-3 (2013).

Early Warning for Infectious Disease Outbreaks: A Q&A with Larry Madoff

The *IQT Quarterly* recently interviewed Larry Madoff, Editor of The Program for Monitoring Emerging Diseases (ProMED-mail). ProMED-mail reports on outbreaks and disease emergence, providing early warning information to a global audience and allowing informed discussions in real time. Madoff discussed ProMED's information dissemination process, the broader community of infectious disease reporting, and the future of outbreak detection and response.

What is ProMED? How did it begin? Who participates?

Founded in 1994, ProMED-mail was intended to harness the Internet in the service of detecting emerging infectious or toxin-mediated diseases, either natural or intentionally caused, that threatened human beings. Its goal is to provide early warning, disseminating information rapidly to a wide audience and allowing informed discussion in real time.

As of 2015, ProMED-mail has 75,000 subscribers in more than 180 countries who receive email reports on outbreaks and disease emergence. Readers can also receive reports via Twitter, Facebook, or an iPhone app. Reports are selected and interpreted by a panel of specialist moderators who provide expert commentary, supply references to previous reports and to the scientific literature, and put the report in perspective for a diverse readership. Reports are simultaneously posted to ProMED's website. ProMED's guiding principles include transparency and a commitment to the unfettered flow of outbreak information, freedom from political constraints, availability to all without cost, commitment to the One Health concept (see below), and service to the global health community.

What kinds of illnesses are reported on ProMED?

ProMED focuses on newly described or unknown diseases, epidemics, and outbreaks, and on the emergence of diseases in new areas or populations. In addition to its focus on human disease, a unique feature of ProMED is its emphasis on the One Health concept, recognizing the importance of diseases that affect plants and animals as well zoonoses (diseases carried by animals that can also infect humans, such as *Escherichia coli* O157, monkeypox, Nipah virus, SARS, and spongiform encephalopathies). About 60 percent

of emerging diseases are zoonoses.[1] Specific examples include ProMED's extensive coverage of the outbreak of foot-and-mouth disease that devastated livestock in the United Kingdom, as well as the recent outbreaks of avian influenza in Europe, Southeast Asia, and the U.S. For this reason, ProMED's staff includes 14 veterinary health specialists with diverse expertise to help us sift through the news of diseases among animals. ProMED provides extensive coverage of less sensational but equally important illnesses, such as dengue fever and norovirus infection (Table 1). Because of their ample coverage in other forums, tuberculosis and HIV infection are not covered by ProMED except in unusual circumstances.

How does it work?

Receipt of information. Each day, ProMED receives reports, many from subscribers, containing new data on outbreaks or diseases, some of which are reported firsthand and some of which are reported from other sources. In addition, our staff searches the Internet and traditional media for relevant items and scans a variety of official and unofficial websites (e.g., national, regional, and local health authorities, and international organizations) looking for recent updates. Since 2007, ProMED has collaborated closely with HealthMap, an organization based at Boston Children's Hospital and Harvard Medical School that automatically scans a large number of news and official websites and continuously searches for infectious disease reports. These reports are provided to ProMED for further analysis and form the basis for some ProMED reports. ProMED's staff of 57 individuals collaborates virtually across 33 countries.

Review and verification. All incoming information is filtered through the "top moderator" — either the editor, or one of the associate editors — who is on duty on a given day. Some reports are rejected immediately because they contain information that is irrelevant, not credible, outdated, or duplicates information contained in previous reports. Most reports are examined carefully and then sent to a member of ProMED's specialty moderators for further review. The panel includes experts in viral diseases, bacterial diseases, plant diseases, veterinary diseases and zoonoses, and epidemiology. Sometimes reports are translated and, on occasion, sent to outside experts for their opinions.

Subsequently, the specialty moderator's main task is to assess the reliability and accuracy of the information. At times, this involves verification of the report from another source, including direct contact with a colleague who might possess firsthand knowledge. In order to help ProMED validate outbreaks of emerging diseases, we established the EpiCore Project (a joint venture between the Skoll Global Threats Fund, HealthMap, ProMED-mail, and TEPHINET) that seeks to maximize the advantage of nontraditional information sources by creating a system for field-based verification of reports from these sources. We are forming a cadre of trained health professionals from around the world that leverages expertise to verify reports received in a geographic proximity through innovative surveillance approaches.

The moderator also edits the piece for content, provides references (both from prior ProMED reports and from the scientific literature), and adds commentary. This commentary is usually brief, with the intention of providing background and perspective. Often multiple reports of the same outbreak or disease entity may be grouped into a single report to enhance clarity and minimize the number of emails our readers receive. Edited reports are returned to the top moderator for final editing, verification, and additional commentary.

Dissemination of information. Finalized reports are simultaneously posted to the ProMED website and distributed to one or more of 19 mailing lists that are based on the interests, language, and region of the subscribers. Approximately one-third of our readers receive the main ProMED-mail list; they receive every report as it is distributed. Other lists are oriented toward animal diseases, such as ProMED-AHEAD (ProMED-Animal Health and Emerging Animal Diseases) or plant diseases. ProMED-EDR (ProMED-Emerging Disease Reports) is designed for readers who want to receive only reports of disease occurrences and do not want to receive discussion, background reports, or announcements. Digest forms of each list are also available. Digest subscribers receive an assemblage of reports approximately once per day. There are also daily and weekly update email lists where subscribers receive a list of post titles and links. These minimize the number of emails but may delay the receipt of a given report.

ProMED is organized into eight regional networks spanning six languages (Arabic, English, French, Portuguese, Russian, and Spanish). This allows information to be tailored by regional concerns, and enhances surveillance in regions where disease

emergence is likely but information resources are less developed.

ProMED's archived database allows users to search 60,000 reports using text, dates, and geographic locations. For example, a user wishing to find reports of Nipah virus in Malaysia could enter these two search terms and receive a list of accessible links. Archives can also be retrieved by email (although without search capability) for those whose Internet connection does not permit Web browser access.

How does ProMED fit into the wider world of infectious disease reporting?

ProMED-mail is a powerful tool, and its growth is testimony to its value. Clearly, however, no single system can detect and report every outbreak of infectious disease worldwide; the need for multiple networks and surveillance systems is widely acknowledged. Other notable systems include:

- The WHO's Global Outbreak Alert and Response Network draws on numerous sources, including its own teams of public health workers, reporting on outbreaks of public health significance and posting some of the information gathered on WHO's website.
- The Global Public Health Intelligence Network, a service of Health Canada, automatically searches the Internet for news stories involving emerging disease threats. However, its use is restricted to a select group of public health officials, and it is not publicly accessible.
- Epi-X, provided by the U.S. Centers for Disease Control and Prevention, is a web-based communications system designed to allow public health professionals (including state and local public health departments) to communicate quickly and securely. It does not seek or allow input from most health practitioners or the general public.
- The Emerging Infections Network is a collaboration between the Infectious Diseases Society of America and public agencies that is designed to allow infectious disease physicians to act as sentinels of disease outbreaks.

Numerous other surveillance systems exist, some directed at specific diseases, regions, populations, or other interest groups. The existence of multiple surveillance systems, official and unofficial, is beneficial from a number of standpoints. The complementary flow of information on the basis of the reporting interests and biases of each network makes it more likely that a given outbreak or emergence of disease will be discovered and reported quickly. Each system serves as an important validation tool for the others. Disease outbreaks that are uncovered by one surveillance system but not by another lead to the recognition of gaps in disease detection. Partial redundancy helps ensure that the overall goal of disease detection is accomplished.

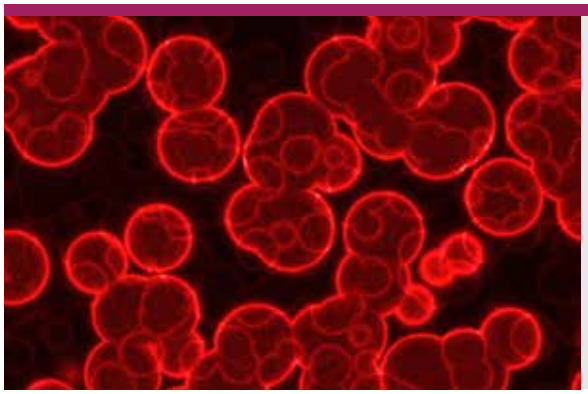
What role will technology play in the future of ProMED and outbreak reporting?

ProMED's focus has never been on technology, but we continue to enhance and refine our operations through collaborations with technically savvy partners such as HealthMap, Metabiota, and EcoHealth Alliance. We are refining our iPhone app and developing an Android version to better allow smartphone users to view our outbreak data and submit information (including photos and geographic coordinates). While many access our reports from our website and social media, we still believe that email is the "killer app" allowing us to push our reports in near real-time to thousands of users. We want our service to be available even to those in remote areas with limited bandwidth or where data is expensive.

We strongly believe that astute human observers are at the heart of global public health and outbreak detection. Technology should therefore be designed to empower individuals to detect and report on unusual occurrences. Smartphones, now more widely deployed, can be one such tool. The ability of concerned and observant people across a variety of disciples to interact virtually — and ProMED itself is a kind of social network — are key to recognizing outbreaks in their earliest stages. Technology should never be prioritized above building human capacity.

What other technology advances do you think are necessary to improve our ability to detect and respond to disease outbreaks?

The ever-widening availability of the Internet, both wired and wireless, is key to our future successes. Other technology directly in the service of astute individuals will include rapid and sophisticated point-of-care (or at least nearby, for example at the district hospital level) diagnostics. In particular, nucleic acid-based technologies will allow specific identification of pathogens. If rapid hemorrhagic fever virus identification had been in the hands of local health



We need to remain broadly vigilant and to develop the human capacity and systems to quickly detect and respond to all types of outbreaks.

care workers in Guinea in late 2013, the response to the Ebola outbreak might have begun sooner, and thousands of cases and deaths prevented. Of course, new therapeutics and vaccines — and rapid ways to develop and study these in the course of an outbreak, are also critical.

What concerns you most as we look at the future of disease outbreaks?

Both private individuals and public officials tend to overreact to disease outbreaks and then become complacent once the immediate threat has waned. There is also a tendency to "fight the last war." For example, there is now a tremendous focus on measures to control Ebola, such as personal protective equipment. But we need to remain broadly vigilant and to develop the human capacity and systems to quickly detect and respond to all types of outbreaks. A robust official public health sector should be complemented by a strong unofficial/NGO sector, including services like ProMED.

Our ability to predict outbreaks is very limited. The next outbreak may be food or waterborne or spread by insects. While we recognize hotspots for disease emergence, they may occur at any time or place. Advances in biotechnology may allow for nefarious development of bioweapons and bioterrorism. The unknown unknowns are my biggest cause of concern.

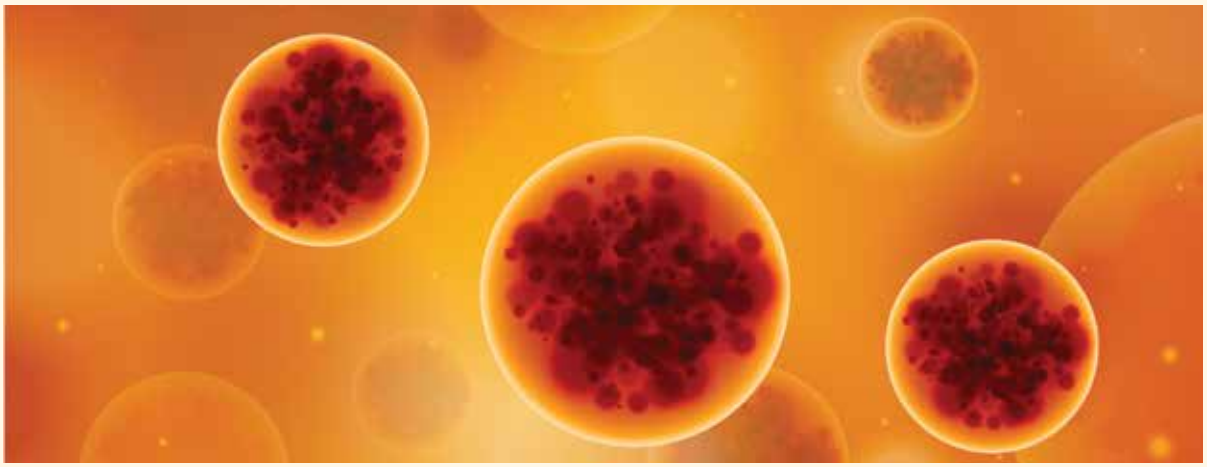
***Lawrence Madoff, M.D.**, is an infectious disease physician specializing in the epidemiology of emerging pathogens, bacterial pathogenesis, and international health. He is Professor of Medicine at the University of Massachusetts Medical School and Lecturer on Medicine at Harvard Medical School. Madoff serves as Director of Epidemiology and Immunization and Deputy State Epidemiologist for the Massachusetts Department of Public Health. Madoff has directed ProMED, the Program for Monitoring Emerging Diseases, since 2002. A graduate of Yale College and Tufts Medical School, he performed his Internal Medicine Residency at New York Hospital-Cornell Medical Center and his Infectious Disease Fellowship at the Harvard Medical School-Longwood program.*

REFERENCE

[1] Woolhouse ME and S. Gowtage-Sequeria, Infect Dis. 11:1842-7.

In Silico Vaccine Design: Accelerating the Response to BioThreats and Emerging Infectious Disease

By Lenny Moise, Sarah Beseme, and Anne S. De Groot



Why Do We Need Vaccines?

Smallpox, polio, measles — control of these lethal diseases is possible because of vaccines. Vaccines are not only the most effective way to eradicate an infectious disease, but are also critically important for protecting first responders and noncombatant (civilian) populations from the consequences of a bioterror event. The U.S. government has expended substantial resources to protect the nation against a potential bioterror event, creating specialized planning and preparedness units within the Departments of State, Defense, Energy, Agriculture, Homeland Security, and Health and Human Services in an effort to comply with the World Health Organization's (WHO) International Health Regulations. These agencies work together to accelerate progress toward a world "safe and secure from infectious disease threats" within the frame of the recent five year Global Health Security Agenda (2014-2018). Several federally subsidized advanced development and manufacturing production facilities in different regions of the country are capable of producing millions of doses of protein-based vaccines. Unfortunately, despite these important advances in the strategic preparedness of U.S. agencies for biodefense, vaccine design remains a significant obstacle to national biodefense. Director of Biomedical Advanced Research and Development Authority (BARDA) Robin Robinson recently stated, "We can produce vaccines faster, but we also need to make vaccines more effective".[1] This is particularly true for the very real threat of new pathogens, for which little is known about the critical antigenic determinants and correlates of immunity, the key parameters used in conventional vaccine design.

Vaccine Payload: the Secret Ingredient

Vaccines harness immunity to prevent disease by training the immune system to recognize and fight infection without requiring exposure to the pathogen. In the 1870s, Louis Pasteur developed the "shake and bake" approach to vaccine protection: growing a pathogen in large flasks, killing it by heat or chemical treatment, or weakening it by producing mutations, and finally injecting the altered whole dead bacteria or virus into a person. Although this method efficiently induces protective immunity, safety concerns about undefined preparations and time scale of development for new pathogens — measured in tens of years — are major challenges that need to be addressed. Killed vaccines require stringent manufacturing processes to assure no live pathogen remains; otherwise, immunization may cause disease. Weakening a pathogen is challenging because of difficulties predicting the necessary extent

of attenuation; over-attenuation results in insufficient protection, while under-attenuation can cause disease. These are serious concerns for approximately 25 percent of the U.S. population, which is immunocompromised and thus would struggle to fight off inadvertent infection.

As knowledge of the immune system and pathogens has expanded, vaccine designers have agreed on three essential components to the modern vaccine construct:

1) **Payload**: pathogen-specific information that allows the immune system to recognize the pathogen;
2) **Adjuvant**: a danger signal that triggers the immune response;
3) **Delivery vehicle**: a packaging vector that carries the vaccine to the right place in the body. While the delivery vehicle and adjuvant may be adapted from one vaccine to the next, identification of the minimum, pathogen-specific payload is a stumbling block which is not addressed by shake and bake methods of the past.

The immune system can respond to pathogen infection through two distinct paths: the innate immune system and the adaptive immune system. Vaccines use properties of the adaptive immune system to protect against infection with the target pathogen. Two main cell types drive the adaptive response: B cells and T cells. Individual B and T cells express a receptor that can recognize small, specific protein sequences, called epitopes. Each of us has a repertoire of hundreds of thousands of receptors, with the potential to recognize hundreds of thousands of epitopes. During infection, pathogen proteins, called antigens, are internalized and broken down into T cell epitopes by specialized cells of the immune system called Antigen Presenting Cells (APC). After transport to the surface of an APC, epitopes are presented to T cells. T cells expressing the matching receptor that recognize the epitope are activated and may proceed to activate B cells. Activated B cells secrete antibodies which target and help clear the pathogen. Activated T cells may also kill cells where pathogens reside (Figure 1). This initial response not only leads to elimination of the pathogen, but also creates B and T cells with "memory" of the pathogen; as a result, when the immune system is re-exposed to the same pathogen, the response is faster and stronger; it clears the pathogen before disease develops. Identification of epitope sequences (i.e., the payload) among as many as thousands of proteins comprising a bacteria or virus is a major challenge for vaccine design.

Computational Solutions for Vaccine Development

Rather than relying on time-honored shake and bake vaccine development approaches, new technologies make it possible to rapidly design vaccines rationally rather than resort to slower trial-and-error discovery of whole pathogen inactivation or attenuation conditions. Over the last two decades, efforts have focused on developing methods and technologies to identify the key proteins that induce an immune response. Computational vaccinology — a.k.a vaccinomics, reverse vaccinology, or genome-derived vaccine design — is driven by the concept that selection and design of the antigen payload is critical for vaccine efficacy. Computational algorithms are developed to identify a minimal set of critical antigens from the genome of a pathogen.

Pathogen-derived epitopes that are presented to T cells, or T cell epitopes, are linear and possess distinct signatures, making it possible for computational algorithms to rapidly scan and identify epitopes in protein sequences. B cell epitopes, in contrast, are considerably more difficult to predict because they are often non-linear or discontinuous in a protein sequence. Generally, antigens that possess many T cell epitopes are strong immunogens. Conversely, antigens that possess few T cell epitopes tend to drive a less effective immune response. Vaccines can also be developed using T cell epitopes, which may provide the minimal, critical information required for protective immunity, even in the absence of the complete antigen. Vaccine design can be improved using bioengineering techniques to include more effector T cell epitopes and conversely delete regulatory epitopes that may suppress immune response. Administration of these engineered immunogens in an appropriate delivery vehicle and adjuvant is the core of a successful vaccine.

Numerous commercial and academic vaccine discovery programs have integrated computational vaccinology tools into the vaccine development process for selecting and optimizing critical antigens. A major milestone for this young field was reached in 2013 when Novartis' Bexsero, a vaccine against meningococcus B, became the first licensed vaccine

developed by a computational approach. Meningococcal vaccines against polysaccharide antigens from *N. meningitidis* were developed for four out of five types of this bacteria (A, C, Y, and W135) but a similar approach for type B was unsuccessful. Its capsular polysaccharide resembles a carbohydrate found on human tissues and prototype vaccines were poorly immunogenic. Using computational tools, a few hundred promising vaccine candidates found on the surface of or exported from the bacterium were rapidly identified from over 2,000 genes. Experimental methods then narrowed the field to three protective antigens.[2] While the entire genome could have been screened experimentally, the computational approach significantly accelerated antigen identification.

Vaccines "On Demand": Accelerated Vaccine Development

The U.S. has spent more than one billion dollars stocking and restocking its anthrax vaccine supply over the past few years — in preparation for just one of 17 possible pathogens considered to be "high risk"

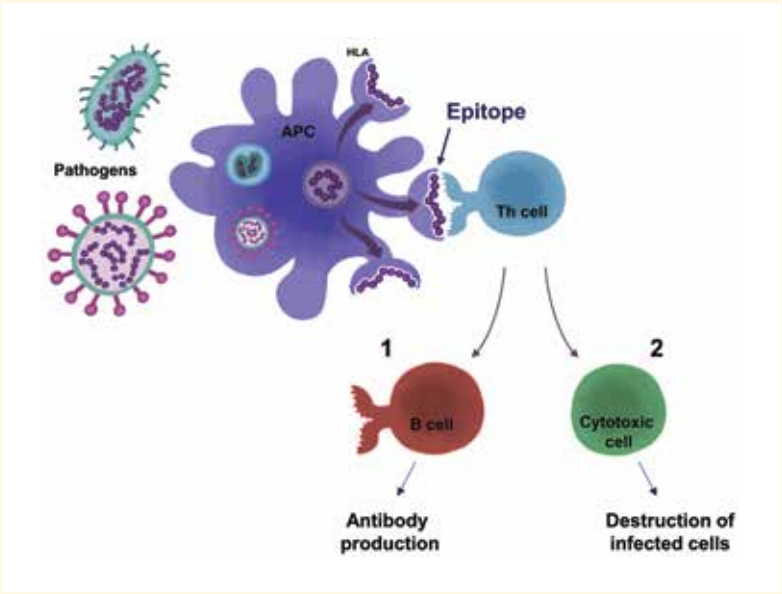


Figure 1 | T cell response to infection by a pathogen. During an infection, pathogens (viruses, bacteria, parasites) are engulfed by antigen presenting cells (APC). Their proteins are broken down into subunits called epitopes and are presented on the surface of the APC through a specific receptor called HLA (Human Leukocyte Antigen). T cells expressing the matching receptor can bind to the HLA-epitope complex, resulting in activation of T cells. Activated T cells induce 1) activation of B cells leading to production of antibodies and 2) activation of cell types able to destroy infected cells.

for a bioterror attack. Is it feasible or even necessary to stockpile a vaccine for every bioterror agent? And what if bioterrorists find a way around stockpiles by engineering a new vaccine-proof version of a pathogen, using well known molecular biology techniques? Making "vaccines on demand" using computational vaccinology, an approach developed by EpiVax, with collaborators at the Vaccine and Immunotherapy Center at Massachusetts General Hospital (MGH), may be the answer to these questions. Funded by the Defense Advanced Research Projects Agency (DARPA) the VaxCelerate Consortium was initiated in 2011. The VaxCelerate program is based on a genome-to-vaccine approach and assumes that as little as one pathogen genome would be available at the start of an outbreak; no prior knowledge of the type of pathogen is required. To meet the time constraints of a realistic scenario, VaxCelerate utilized a distributed R&D model. The network included experts in the fields of protein engineering, expression and purification (Pfenex), integrated computational epitope prediction and vaccine construction (EpiVax), peptide synthesis/design (21st Century Biochemicals), Good Manufacturing Practice (GMP)/Good Laboratory Practice (GLP) and testing standards (MGH and MPI Research). The combined adjuvant and delivery vehicle for this vaccine form a self-assembling vaccine platform developed by MGH. The payload is developed by EpiVax's computational algorithms that analyze the genome of interest to identify candidate vaccine antigens and map their T cell epitopes. Epitopes predicted to be immunogenic and broadly reactive are selected. Multi-epitope peptides are designed using EpiVax's vaccine building tools and synthesized in less than 10 days. Assembly of the vaccine is performed by mixing peptides and the delivery vehicle platform in saline, which can be performed on-site at the time of vaccination (Figure 2).

In a first realistic scenario test, the consortium demonstrated that rapid, scalable vaccine generation against Lassa Fever could be accomplished in less than 120 days (VaxCelerate-2).[3] T cell dependent immune responses were observed in transgenic mice immunized by the vaccine, demonstrating efficacy. Following successful completion of the DARPA contract, the Defense Threat Reduction Agency (DTRA) contracted the VaxCelerate Consortium to develop a Q fever vaccine (Q-VaxCelerate).

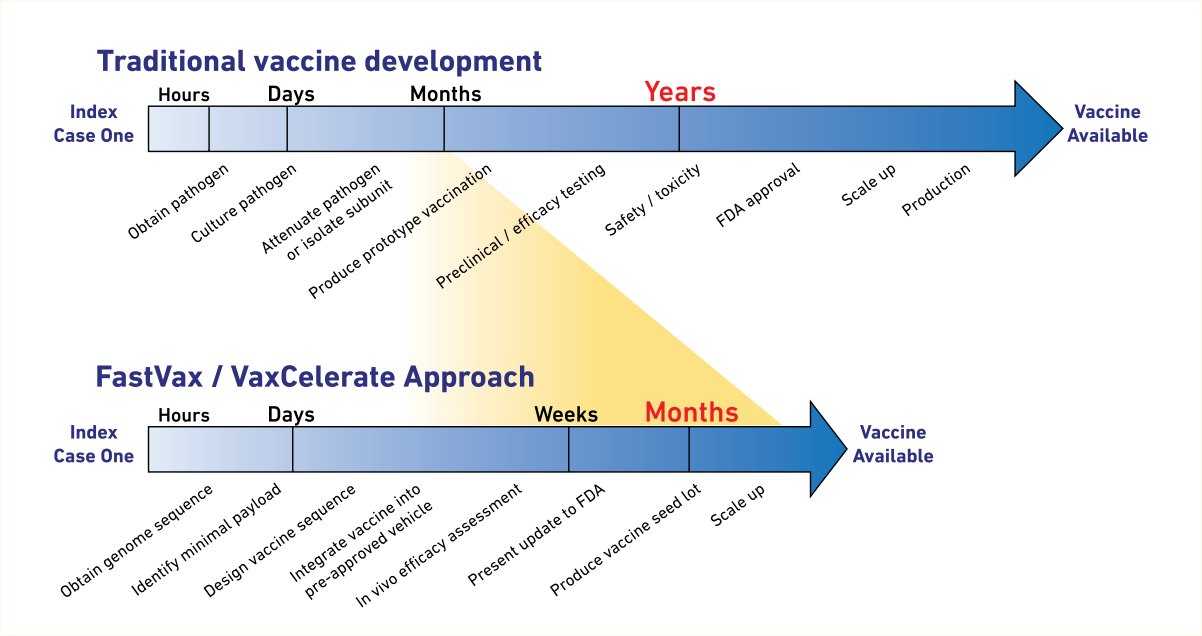


Figure 2 | Accelerated vaccine design with the Vaxcelerate program. Using a coordinated approach, the Vaxcelerate program enables the production of a vaccine within months of an outbreak.

Evaluation of Existing Vaccines: Toward Better Flu Vaccines

Computational vaccinology is not only a powerful tool in the fight against emerging pathogens that threaten human health in a global emergency scenario; it can also be used to combat seasonal viruses such as influenza. *In silico* vaccine design tools can be used to evaluate the efficacy of existing vaccines against new pathogens, or to predict the efficacy of a new vaccine against an emerging pathogen. Three examples are given in the next section.

Predicting the need for a new vaccine in case of a new outbreak strain

When novel swine-origin influenza virus (H1N1) emerged in Mexico and the Western U.S. in March 2009, the U.S. Centers for Disease Control and Prevention (CDC) and WHO predicted widespread transmission and high mortality because antibodies to seasonal influenza were not protective against the new viral strain. Prior to development of a new H1N1 vaccine, EpiVax performed a computational comparison of the T cell epitope content of the new 2009 H1N1 viral strain and the 2008–2009 seasonal flu vaccine. EpiVax predicted cross-reactivity of T cell epitopes, suggesting that prior influenza exposure or seasonal influenza vaccination might confer protection against morbidity. The predicted cross-reactivity was later confirmed using samples from H1N1 exposed subjects and 2008–2009 seasonal flu vaccine immunized subjects. Cross-sectional studies of hospitalized patients and challenge studies in ferrets and mice confirmed the hypothesis that seasonal H1N1 exposure or vaccination could protect against severe pandemic H1N1 disease.[4, 5, 6]

Predicting vaccine efficacy

A new avian H7N9 flu virus emerged in humans in China in 2013. The virus has been associated with high mortality in humans. The reservoir for this virus has not been identified, although closure of live poultry markets had a dampening effect on human cases. The seasonal recurrence of H7N9 outbreaks in China and the potential for the virus to be efficiently transmitted from human to human poses a serious threat to public health. Unfortunately, attempts to design an effective vaccine have been largely unsuccessful. Whole-inactivated H7N9 vaccines were poorly immunogenic and failed to protect ferrets from infection. Subunit vaccines developed using the viral protein H7 Hemagglutinin (HA), a widely used antigen, have also been poorly immunogenic in humans.[7, 8] EpiVax predicted the poor immunogenicity

of the H7N9 virus in an analysis within 24 hours of sequence publication on the Global Initiative on Sharing Avian Influenza Data website. Not only did the H7N9 virus genome contain an extremely low number of T cell epitopes, but the HA and PB1 antigens contained epitopes that were highly cross-conserved with the human genome. Previous studies have demonstrated that persistent viruses such as Hepatitis C Virus or HIV and other pathogens escape the immune response using epitopes with a high degree of cross-conservation with the human genome.[9, 10] This mechanism is called immune camouflage, as these epitopes activate a sub-group of T cells (called regulatory T cells) that dampen the immune response, rather than activate it. Assays measuring the immune response in human T cells after exposure to H7N9 HA and PB1 epitopes confirmed the observation.[11] Immunophenotyping data demonstrate a strong link between these "human-like" epitopes and regulatory T cell expansion, providing a possible explanation for the poor immunogenicity observed with whole inactivated H7N9 and recombinant H7 HA vaccines.

Improving vaccine safety

Although rare, in some instances, vaccination has been associated with auto-immune diseases. Examples include the onset of Guillain-Barré Syndrome following immunization with the influenza vaccine in 1976.[12] More recently, an increase in the number of cases of narcolepsy was reported following exposure to and vaccination with 2009-2010 H1N1 influenza vaccine in China and Europe.[13, 14] Homology between a peptide derived from an influenza protein and a portion of hypocretin receptor 2, a protein involved in narcolepsy, was identified by computational tools. New epitope homology tools developed at EpiVax, such as the JanusMatrix algorithm, may be useful for uncovering similar relationships between vaccine antigens and self proteins; further studies are necessary to determine whether such cross-reactivity is predictive of adverse effects.[15]

In summary, integration of new computational tools into the vaccine design process makes it possible to design vaccines in the shortest time possible once the DNA sequence from an emerging infectious disease or biowarfare pathogen is available. Multi-disciplinary platforms that offer an efficient way to quickly produce millions of doses of a specific vaccine are now possible, enabling nations to rapidly respond to outbreaks of both known and novel pathogens.

***Anne S. De Groot, M.D.**, earned her B.S. at Smith College and her M.D. at University of Chicago's Pritzker School of Medicine. After being awarded her first NIH grant, she developed cutting-edge tools for computational vaccinology at Brown University. With an award from the Slater Biotechnology fund, she co-founded EpiVax, Inc. with Bill Martin and has served as CEO/CSO and President of the company since its inception. De Groot established the Institute for Immunology and Informatics (iCubed) at the University of Rhode Island as a center of excellence in Computational Vaccinology. In 2015, she was invited to testify at the Blue Ribbon Panel on Biodefense by Former Senator Lieberman and Governor Ridge. She was recognized as one of the 50 most influential people in vaccinology in 2014. Her main areas of research interest are: Computational vaccinology, "Vaccines on Demand" for biodefense, development of personalized vaccines, and the dance between immunogenicity and tolerance.*

***Lenny Moise, Ph.D.**, is Director of Vaccine Research at EpiVax, Inc. Dr. Moise received his Ph.D. from the Department of Molecular and Cellular Biology and Biochemistry at Brown University in Providence, RI. He joined Dr. Anne De Groot's laboratory at Brown University to study T-cell epitope-driven vaccination and protein therapeutic immunogenicity. At EpiVax, Inc. Dr. Moise leads T-cell epitope-driven vaccine development projects using a genomes-to-vaccine approach that combines cutting edge immunoinformatic and immunologic methods. He is a founding faculty member of the URI Institute for Immunology and Informatics (iCubed).*

***Sarah Beseme, Ph.D.**, earned her Ph.D. at the University of Lyon, France. During her postdoctoral fellowship at Brown University, Providence RI, she worked on the discovery of new therapeutics for the treatment of liver cancer. Her involvement in clinical studies with Hepatitis B and C patients led her to the field of infectious diseases and she started a collaboration with Annie De Groot to implement HPV vaccination in Mali, Western Africa. Dr Beseme works at EpiVax, Inc. as part of the scientific team to develop new vaccine and safer protein therapeutics.*

ACKNOWLEDGMENTS

The authors thank Steven Vessella for his careful review of the manuscript and his assistance for the design of the figures.

BIBLIOGRAPHY

1 De Groot AS, Einck L, Moise L, Chambers M, Ballantyne J, Malone RW, Ardito M, Martin W. Making vaccines "on demand": a potential solution for emerging pathogens and biodefense? Hum Vaccin Immunother. 2013 Sep;9(9):1877-84

2 Pizza M, Scarlato V, Masignani V, Giuliani MM, Aricò B, Comanducci M, Jennings GT, Baldi L, Bartolini E, Capecchi B, Galeotti CL, Luzzi E, Manetti R, Marchetti E, Mora M, Nuti S, Ratti G, Santini L, Savino S, Scarselli M, Storni E, Zuo P, Broeker M, Hundt E, Knapp B, Blair E, Mason T, Tettelin H, Hood DW, Jeffries AC, Saunders NJ, Granoff DM, Venter JC, Moxon ER, Grandi G, Rappuoli R. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science. 2000; 287(5459):1816-20.

3 Leblanc P, Moise L, Luza C, Chantaralawan K, Lezeau L, Yuan J, Field M, Richer D, Boyle C, Martin WD, Fishman JB, Berg EA, Baker D, Zeigler B, Mais DE, Taylor W, Coleman R, Warren HS, Gelfand JA, De Groot AS, Brauns T, Poznansky MC. VaxCelerate II: rapid development of a self-assembling vaccine for Lassa fever. Hum Vaccin Immunother. 2014;10(10):3022-38

4 Hancock K, Veguilla V, Lu X, Zhong W, Butler EN, Sun H, Liu F, Dong L, DeVos JR, Gargiullo PM, Brammer TL, Cox NJ, Tumpey TM, Katz JM. Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. N Engl J Med 2009; 361:1945-52

5 Ellebedy AH, Fabrizio TP, Kayali G, Oguin TH 3rd, Brown SA, Rehg J, Thomas PG, Webby RJ. Contemporary seasonal influenza A (H1N1) virus infection primes for a more robust response to split inactivated pandemic influenza A (H1N1) Virus vaccination in ferrets. Clin Vaccine Immunol. 2010 Dec;17(12):1998-2006

6 Price GE, Soboleski MR, Lo CY, Misplon JA, Pappas C, Houser KV, Tumpey TM, Epstein SL. Vaccination focusing immunity on conserved antigens protects mice and ferrets against virulent H1N1 and H5N1 influenza A viruses. Vaccine 2009; 27:6512-21

7 Couch RB, Patel SM, Wade-Bowers CL, Niño D. A randomized clinical trial of an inactivated avian influenza A (H7N7) vaccine. PLoS One 2012; 7:e49704

8 Wong SS, Jeevan T, Kercher L, Yoon SW, Petkova AM, Crumpton JC, Franks J, Debeauchamp J, Rubrum A, Seiler P, Krauss S, Webster R, Webby RJ. A single dose of whole inactivated H7N9 influenza vaccine confers protection from severe disease but not infection in ferrets. Vaccine. 2014 Jul 31;32(35):4571-7

9 Losikoff PT, Mishra S, Terry F, Gutierrez A, Ardito MT, Fast L, Nevola M, Martin WD, Bailey-Kellogg C, De Groot AS, Gregory SH. HCV epitope, homologous to multiple human protein sequences, induces a regulatory T cell response in infected patients. J Hepatol. 2015 Jan;62(1):48-55

10 He L, De Groot AS, Gutierrez AH, Martin WD, Moise L, Bailey-Kellogg C. Integrated assessment of predicted MHC binding and cross-conservation with self reveals patterns of viral camouflage. BMC Bioinformatics. 2014;15 Suppl 4:S1

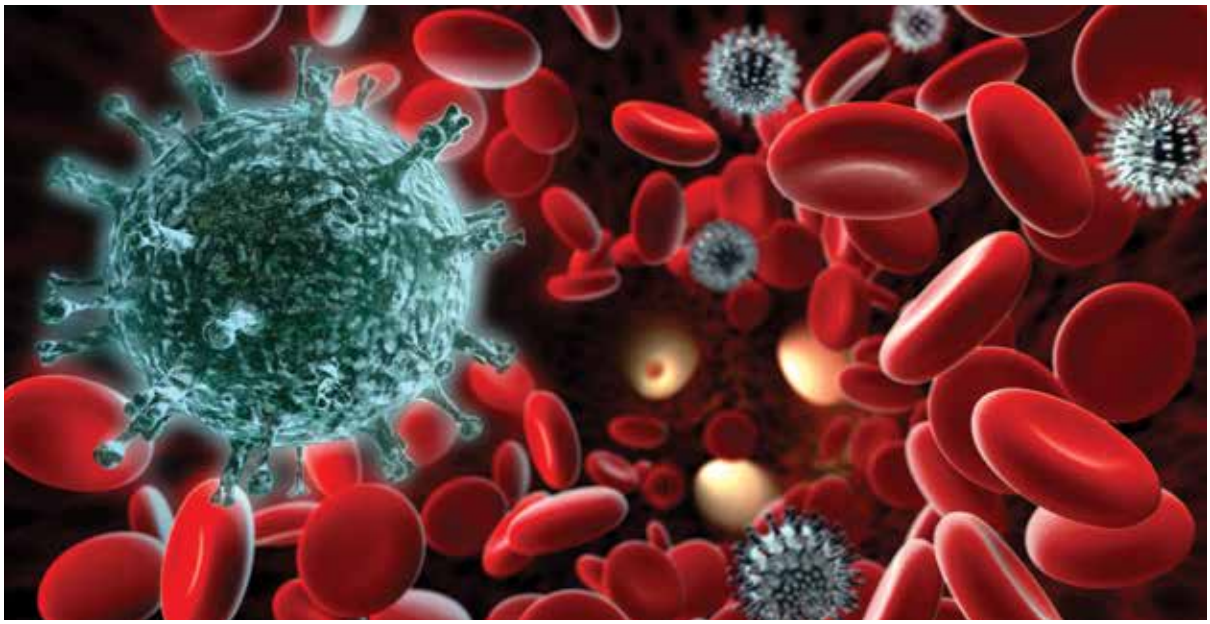
11 Liu R, Moise L, Tassone R, Gutierrez AH, Terry FE, Sangare K, Ardito MT, Martin WD, De Groot AS. H7N9 T-cell epitopes that mimic human sequences are less immunogenic and may induce Treg-mediated tolerance. Hum Vaccin Immunother. 2015 Sep 2;11(9):2241-52

12 Marks JS, Halpin TJ. Guillain-Barré syndrome in recipients of A/New Jersey influenza vaccine. JAMA 1980;243(24):2490-4

13 Wijnans L, Lecomte C, de Vries C, Weibel D, Sammon C, Hviid A, Svanström H, Mølgaard-Nielsen D, Heijbel H, Dahlström LA, Hallgren J, Sparen P, Jennum P, Mosseveld M, Schuemie M, van der Maas N, Partinen M, Romio S, Trotta F, Santuccio C, Menna A, Plazzi G, Moghadam KK, Ferro S, Lammers GJ, Overeem S, Johansen K, Kramarz P, Bonhoeffer J, Sturkenboom MC.. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine 2013;31(8):1246-54

14 Han F, Lin L, Warby SC, Faraco J, Li J, Dong SX, An P, Zhao L, Wang LH, Li QY, Yan H, Gao ZC, Yuan Y, Strohl KP, Mignot E.. Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. Ann Neurol 2011;70(3):410-7

15 Moise L, Gutierrez AH, Bailey-Kellogg C, Terry F, Leng Q, Abdel Hady KM, VerBerkmoes NC, Sztein MB, Losikoff PT, Martin WD, Rothman AL, De Groot AS. The two-faced T cell epitope: examining the host-microbe interface with JanusMatrix. Hum Vaccin Immunother. 2013 Jul;9(7):1577-86



The Weakest Link in Diagnostics May Be the Sample Itself

By Mark D. Lim

Anyone who has had blood drawn at a doctor's office is familiar with much of the information that results from its analysis, such as glucose, cholesterol, lipids, and cell counts. In the past two decades, two trends have given rise to a revolution in the next generation of diagnostic technologies.

First, molecular biology is increasingly finding molecules in bodily fluids (such as blood and saliva) and tissues that provide information about a patient's health state to guide the decisions made by a physician ("biomarkers"). Second, engineering advances have enabled these tests to be performed with equipment that is smaller, less expensive, and requires less sample for analysis. These trends are creating market opportunities that, for DNA sequencing technologies alone, are predicted to exceed $20 billion with opportunities beyond niche diagnostic markets.[1] This is encouraging more entrepreneurs to join this rapidly developing market by applying advanced analytics and novel sensors to increase the speed, precision, and accuracy for interrogating human samples, while also reducing per-analysis costs.

The same advances in diagnostic technology are finding use outside of the clinical laboratory. Less expensive, more powerful diagnostics allow public health researchers to screen larger groups of people to determine if a population is at-risk of developing or harboring disease. Such broad applications of genomic- and proteomic-based diagnostics to public health are as vulnerable as any research to flaws that prevent their reproducibility, as has been reported recently.[2] While there are many factors that may contribute to irreproducibility in the use of diagnostic technology, there are some sources of data variability and bias that can be mitigated when engineering the workflow from human sample to molecular answers on a population scale.

The Problem of Variability in Molecular Analysis

A central task for most molecular diagnostics is the measurement of a biomarker or panel of biomarkers; that is, determining which molecules (typically proteins, DNA, RNA, or small molecules) are present and in what quantities, and how variations in the presence and/or quantity of those molecules correlate with

health or illness. For diagnostic technologies that are being developed to distinguish the well from the sick, track the progress of disease outbreaks, or predict susceptibility to future illness, the characteristics of the analytes detected by a platform must be in context of the entire population intended to be screened. This statement seems obvious, but several efforts to develop diagnostics fail to account for regional normality; that is, samples used to create and benchmark the test did not represent the actual population where it will be used. Epidemiologically-relevant demographics that developers must consider include age, gender, natural history of disease, lifestyle, co-infections, geographies, and environmental exposures. Failure to take these variables into account can invalidate a given diagnostic technology or biomarker for application to a specific population or geography.

The biomarker(s) targeted by a diagnostic also need to be chosen with care — and it's important to consider their quantitative abundance in any given volume of fluid.[3] For example, the gold standard for determining the concentration of glucose in blood is analysis of glucose itself. However, there is significant market demand in diabetes care for methods that can determine blood glucose without the need for a finger stick or or venipuncture, and the R&D landscape is crowded with research aiming to measure glucose in saliva or tears. However, any solution that is acceptable to physicians and regulators (giving good metrics for patient care) and patients (non-invasive and convenient) still requires additional foundational research to benchmark glucose measurements in non-invasive specimens to "gold-standard" blood-based measurements.[4]

Pre-Analytical Variability Threatens the Reliability of Many New Technologies: Sample Collection

In addition to the inherent variability among human subject populations, variability in the handling of sample materials themselves can render study results irreproducible.[5] The high analytical sensitivity of next-generation diagnostics potentially amplifies these sources of sample-to-sample variation. When this variability noise is confused with signal, it confounds the very purpose of disease surveillance and screening; what appears to be real person-to-person differences can be the result of mere differences in sample collection and handling. Ignoring variability in the collection and handling of samples can result in the detection of a signal that should have been noise (a false positive result), which consumes resources to verify. Conversely, a signal that was buried in the noise results in the failure to detect a true risk (a false negative result).

Sampled fluids and tissues begin to change as soon as they are removed from the human body, being broken down by both internal factors (e.g., degradative enzymes) and external (e.g., the presence of microorganisms, in samples like saliva and stool). When developing a reliable diagnostic technology, researchers and end users must account for these sources of instability that might result in unacceptably large sample-to-sample variability. Additional variability can also be introduced by inconsistent pre-analytical methods for collecting, transporting, and preparing a sample. These processes are often manual and vary in waiting times (samples left on benchtops before further processing), temperatures, humidity, reagents, and handling conditions.

The impact of sample collection conditions on the reproducibility of diagnostic results is greater in resource-limited environments, particularly for unstable analytes such as RNA or proteins. As mentioned above, samples begin to degrade as soon as they are collected, and refrigeration or freezing are typical means of slowing sample degradation. Preservatives are often added to samples as well, and must be chosen carefully. For example, preservatives added to blood storage tubes (such as anti-coagulants, clotting agents, etc.) have been shown to interfere with later analysis by various platforms.[5, 6]

Consistency of sample preparation is another goal of diagnostic technology developers. All diagnostic analytical platforms require the target molecules to be purified and concentrated from a sample before analysis. This process, called sample preparation, is often as, if not more, complex than the analysis itself. On average two-thirds of the real estate and fluidic components on commercially-available test cartridges are dedicated to sample preparation, typically customized to the analytic platform, biospecimen type, and specific analyte(s).

Developers of new diagnostic technology and its end-to-end workflow therefore must, as an inherent part of their process, evaluate all potential sources of variability for tests that analyze molecules as diverse as nucleic acids (DNA and RNA), proteins, and small molecule metabolites. Unfortunately, there is no "cure-all" method to mitigate all sources of pre-analytical variability for all platforms and analytes. Technology developers must try their best to standardize the means of collecting and transporting samples to the instrument. A consistent, effective, and rigorously followed protocol for sample

collection and preparation is essential for mitigating post-collection alterations caused by the illustrative factors mentioned above.

To illustrate the process of developing a sample collection technology, we discuss below one example, among the simplest and most useful methods available in resource-constrained parts of the world.

Case Study: The Promise of Dried Blood Spot Cards

Dried blood spot cards continue to capture the imagination of multiple communities since they were first demonstrated by Robert Guthrie to simplify newborn screening in the early 1960s; DBS is still central for routine screening of specific metabolic, endocrine, and genetic disorders within the developed world. For this use case, DBS offers the ability to take a small volume of blood (15 to 60 µL) from a newborn without a needle (beyond a lancet for the heel prick) and place it onto an absorbent card. DBS is also useful in the drug development setting, as researchers use them to collect small volumes of blood from small, fidgety animals while maximizing the number of sampling opportunities. DBS cards are also useful in public health and population-wide disease surveillance because they simplify the logistics of transporting a self- or simply-collected sample from a remote location to a central laboratory. DBS cards are inexpensive to manufacture by robust roll-to-roll processes, making them an effective technology for large population screening.

Because of their usefulness, the biopharmaceutical industry has devoted significant resources to evaluate the performance of DBS. Most studies have uncovered multiple ways in which DBS introduce variability into downstream analyses. Commonly cited issues include:

- Patient-to-patient differences in red blood cell concentrations can affect the measured levels of other blood components in an unpredictable manner;
- DBS cards themselves pre-separate blood components in a variable manner before drying, in a process called chromatographic separation; and
- Efficiencies of extracting blood components from cards varies with the target component, card type, solvent, and other experimental conditions.

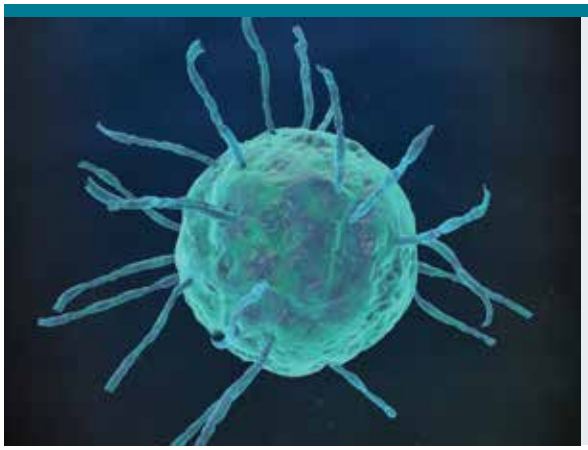
When collecting samples from newborns, or in drug development, DBS is typically used in controlled settings like a hospital or centralized laboratory. This is not the case, however, when DBS is used as the sample collection medium for diagnosis or disease surveillance in remote regions. In such areas, variability in later analysis of DBS-collected samples may be caused by:

- Contamination by the handler pre- or post-collection
- Contamination from dust, dirt, or insects when dried on an open surface
- Contamination from contact with other non-dried cards
- Variation in drying times due to temperature and humidity
- Degradation of analytes from exposure to sunlight
- Fluctuations in environmental conditions after drying, and during storage and transportation

There have been far fewer studies evaluating DBS for use in remote diagnosis or surveillance. Those studies that exist similarly reveal unpredictable performance, highlighting the need to carefully characterize any potential variability in relation to specific analytes and platform technology. For example, DBS is often recommended when collecting samples from patients in remote areas when assessing HIV viral load to monitor efficacy of a specific therapy, or perform early infant diagnosis. However, the ability of DBS to permit uniform extraction and processing of RNA (the key analyte in this case) has only been reported recently. Some researchers report inconsistent measurements of HIV viral load that appear to be dependent on the chosen analytical platform.[7] Under these circumstances, a quantified measurement that incorrectly represents an individual's health status can result in an incorrect assessment of the patient's response to anti-HIV therapy.

No approach has been shown to mitigate all sources of variability when using DBS as the front-end sample, collection, and transport medium. In fact, DBS cards are largely unchanged since initially described by Dr. Guthrie. Non-cellulose DBS cards have been developed for cases where cellulose may affect the extraction of target molecules. Other versions of DBS cards include preservatives that stabilize nucleic acids and proteins. DBS-like formats have also been developed to collect and store derivatives of whole blood that are prepared on- or off-card, such as dried plasma and dried serum cards. There are also DBS accessories that protect cards from environmental conditions or cross-contamination, enhance dehydration through desiccants, and devices that control the volume deposited onto the card.

As is true for any sample collection technology, the value of DBS in any strategy — research, diagnostics, or surveillance — needs to be counterbalanced with the resources required to evaluate and minimize sources of data variability. This can only be done by identifying each step in the integrated collection-to-result workflow and evaluating where variable processes and



To fully exploit the remarkable advances in diagnostic science and engineering that are now emerging, it is essential to account, to the greatest extent possible, for the experimental variability that exists before the actual diagnostic analysis.

conditions impact downstream analytical results. An important set of basic principles was published through a collaboration between multiple pharmaceutical companies and the FDA.[7,8] Even though these procedures are focused on a drug development use case, they serve as an important resource to estimate the level of rigor for qualifying DBS processes and technologies for other use-cases.

Hope Comes From the Intersection of Technology and Rigorous Processes

To fully exploit the remarkable advances in diagnostic science and engineering that are now emerging, it is essential to account, to the greatest extent possible, for the experimental variability that exists before the actual diagnostic analysis. Deep diligence in all these variables — disease, epidemiology, validation of biomarkers and the sample type from which they are isolated, as well as considerations for sample collection, preservation and transport, and the choice of a compatible analytical platform — is laborious, but essential to maximize the value of data in disease surveillance and public health assessments. This diligence is best achieved by analyzing replicate samples collected and handled under the range of concentrations and conditions that simulate the real-world user setting, with an eye towards assessing impact on analytical linearity, trueness, detection limits, consistency, and precision. When done rigorously, this process leads to an end-to-end, sample-to-answer diagnostic architecture that can provide value across the entire public health spectrum, from health monitoring to the surveillance and eradication of endemic disease, and the quenching of disease outbreaks. Q

Mark D. Lim *is a Program Officer at the Bill and Melinda Gates Foundation's diagnostics team, focused on investing in the development of diagnostics for neglected tropical diseases and cervical cancer. Prior to joining the Gates Foundation, Lim served as an Associate Director at FasterCures, a Center of the Milken Institute, focused on mechanisms for multi-stakeholder partnerships to biomedical research, in addition to providing counsel to several patient organizations with active funding programs. Previously, Lim served a supporting role as the Chief of Technical Staff to Col. Daniel Wattendorf at DARPA's Defense Sciences Office, on a program that had efforts in new diagnostic and vaccine capabilities. Lim received his Ph.D. in Chemistry at the University of California, Santa Barbara and completed an NCI-funded postdoctoral fellowship in cancer nanotechnology at the University of California, San Francisco and the University of California, Berkeley.*

REFERENCES

1. DNA Sequencing Market Will Exceed $20 Billion, Says Illumina CEO Jay Flatley, Forbes, April 2015. http://www.forbes.com/sites/33lu ketimmerman/2015/04/29/qa-with-jay-flatley-ceo-of-illumina-the-genomics-company-pursuing-a-20b-market, accessed on Oct 2015.
2. Trouble at the lab, The Economist, October 2013, http://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble, accessed on Oct 2015
3. Bond MM, et al, Drop-to-drop variation in the cellular components of fingerprick blood, Am. J. Clin. Pathol., 2015, 144, 885-894.
4. Mascarenhas P, et al, Effect of Diabetes Mellitus Type 2 on Salivary Glucose – A Systematic Review and Meta-Analysis of Observational Studies, PLoS One, 2014, 9(7), e101706.
5. Lim MD, et al, Before you analyze a human specimen, think quality, variability, and bias, Anal Chem, 2011, 83(1), 8-13.
6. Yin P, et al, Preanalytical aspects and sample quality assessment in metabolomics studies of human blood, Clin. Chem., 2013, 59(5), 833-45.
7. Sawadogo S, et al, Limited Utility of Dried-Blood- and Plasma Spot-Based Screening for Antiretroviral Treatment Failure with Cobas Ampliprep/TaqMan HIV-1 Version 2.0, J Clin Microbiol, 2014, 52(11), 3878-83.
8. Bowen CL, et al, Investigations into the environmental conditions experienced during ambient sample transport: impact to dried blood spot sample shipments, Bioanalysis, 2011, 3(14), 1625-33.

To supplement the ***IQT Quarterly***'s focus on technology trends, *Tech Corner* provides a practitioner's point of view of a current challenge in the field and insight into an effective response.

A technology overview from IQT portfolio company Quanterix

Medical laboratory analysis is moving steadily and quickly into molecular biology, in which tests can assess health based on DNA sequence, or on changes in the amount of certain proteins in blood. There are tests currently available to detect the levels of over 100 different proteins in human blood, changes in which have been correlated with changes in health or disease. Nearly all of these tests are called immunoassays because they rely on the use of immune system proteins called antibodies as the essential ingredient in the test.

Figure 1 | Target proteins available in blood. Of the proteins in blood whose concentrations are high enough (millimolar to picomolar) to detect by currently available means, 171 are the targets of FDA-approved diagnostic tests. At least another 4,000 proteins are present in blood whose levels, were they detectable, might be useful indicators of disease.

Antibodies have the property of binding tightly to a specific molecular target (and only that target), giving immunoassay tests the specificity required to detect target proteins in a sea of proteins.

The most commonly used immunoassay is called ELISA (Enzyme-Linked Immuno-Sorbent Assay), which is used by medical labs and researchers to detect specific proteins of interest in liquid samples. The human genome encodes about 25,000 proteins. Of these proteins, no more than 10 percent are in sufficient concentration to be reliably measured with conventional ELISA. What clinical insights lie within the other 90 percent? Quanterix's revolutionary Simoa technology unlocks a world of insight into disease detection, diagnosis, and patient treatment while meeting the demands of today's laboratory (Figure 1).[1,2]

Single Molecule Array Technology

Quanterix's technology reveals what lies beneath the water level in Figure 1 by increasing the sensitivity of ELISA by, on average, 1,000-fold. In ELISA, the presence of a target protein generates a color or fluorescence in a small amount of liquid in a plastic laboratory dish. Quanterix's improvement on ELISA is both simple and profound: Simoa performs ELISA in volumes that are approximately 2 billion-fold smaller than those now used in laboratories. In doing so, the colored signal created by the presence of target is confined in a tiny space, and the local concentration of color, even that created by a single target molecule, is very high. For this reason, we call our method Simoa (Single Molecule Arrays). Quanterix

fully automated the Simoa assay platform in its first instrument, the Simoa HD-1 Analyzer.

Unparalleled Sensitivity

In conventional immunoassays, reaction volumes are relatively large (100 µL, or one-tenth of a cubic milliliter) and it takes millions of molecules generating billions of fluorophores generated by the enzyme substrate complex diffusing in this dilute solution before an optical signal can be obtained. In contrast, the signal generation volume in a Simoa assay is 2 billion times smaller, so a single target molecule in a sealed microwell quickly generates enough fluorophores to be measured using conventional fluorescence imaging. When concentrations of the target analyte reach levels above which digital calculations are meaningful, the system's proprietary algorithm converts to an analog measurement, ensuring accuracy across a wide dynamic range (greater than four logs).

Robust Multiplexing

The optical system of the HD-1 Analyzer can detect a range of colors. By incorporating a range of colored dyes into the beads used to perform Simoa assays, samples can be tested for the presence of up to 10 different proteins simultaneously with no loss of precision across a broad dynamic range. This practice, called multiplexing, preserves samples by allowing multiple results from small volumes, saves costs on consumables, and can dramatically increase throughput for larger experiments.[3]

Complete Automation

The Simoa HD-1 Analyzer performs the biochemical manipulations of the assay with an on-board robot, detects the bead-bound targets, and calculates the concentration of each target protein. Users load tubes or plates containing samples, assay reagents, and proprietary consumables, select the appropriate protocol on an intuitive touch screen interface, and read results on the instrument, or export data to commonly used software packages and Laboratory Information Management Systems (LIMS).

A Wide Spectrum of Applications

Working in collaboration with Quanterix, several research groups have already made and published provocative findings using Simoa.

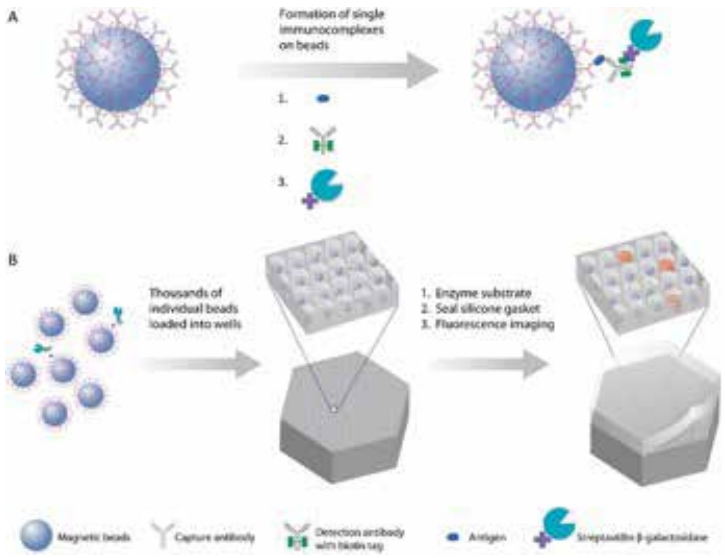


Figure 2 | Diagram of a single molecule immunoassay.

A - Beads coated with capture antibodies are the solid platform or the building the ELISA sandwich.

B - Beads with complete reactions are trapped in femtoliter-scale wells and exposed to substrate. If even a single sandwich is present on a bead, a well will appear fluorescent in the field of an optical microscope.

Life Science Research: Quanterix has developed a robust menu of ultrasensitive assays for proteins of significant interest across the medical and life science research community, including targets for neurology, oncology, inflammation, and cardiac and infectious disease. In addition, Simoa's open platform and homebrew kit allow researchers in the biopharmaceutical industry, academia, and other laboratories to easily develop their own ultrasensitive assays to proteins of interest. By partnering with the diagnostics company bioMerieux, Quanterix has also ensured a product continuum for the Simoa technology, from research use only (RUO) to the in vitro diagnostic (IVD) markets.

Neurology: Quanterix has developed a strategic focus in neurology and neurodegeneration and is working with a rapidly growing network of academic researchers and pharmaceutical and biotech partners. This work has already been published in several pivotal publications and been recognized by funding from several

organizations. Key areas so far are concussions/traumatic brain injury, Alzheimer's disease, and Parkinson's. The National Football League has awarded Quanterix significant grants to further explore Simoa for its potential to diagnose concussions with a simple blood test, which could revolutionize the way athletes are treated when they have a suspected head injury.

Oncology: The ability to detect oncogenic related biomarkers at ultra-low levels is enabling new options for diagnostics and treatment in oncology. Simoa based ultra-sensitive biomarker assays can be used to monitor cancer risk, identify early stage cancers, and discriminate between benign and malignant cells. Simoa based biomarkers can be used to predict disease outcomes, predict progression free survival and monitor reoccurrence. Additionally, these assays can be used to monitor sensitivity to therapy and to aid in treatment decisions.

Inflammatory Disease: Conventional immunoassays often lack the sensitivity required to measure many of the inflammatory cytokines. This lack of sensitivity has limited the ability of clinicians to assess disease severity based on the levels of these critical biomarkers or to monitor therapeutic responses. The ability to detect these markers in normal healthy patients, in which the levels are often quite low, is also a limiting factor in many studies.

Infectious Disease: Infectious disease is one of the hottest areas of research today and researchers continue to seek ways to detect and diagnose infections earlier and more accurately. In a ground-breaking paper published in the *Journal of Virological Methods,* researchers were able to demonstrate that Quanterix's Simoa technology provided a 3,000-fold improvement in sensitivity over conventional immunoassays for identifying acute HIV infection, a level of sensitivity equivalent to the gold standard of nucleic acid testing (such as PCR), but at much lower cost.[4] A lower-cost HIV assay can make testing more accessible to people in resource-constrained countries where HIV prevalence is high. Furthermore, detection early during infection is key to controlling the spread of HIV, since recently infected persons are 10 times more infectious than persons who have developed an immune response to the virus. Until now, only nucleic acid-based tests were sensitive enough to detect HIV at this critical stage. With Simoa, however, early testing can be performed at scale, for lower costs, making early stage detection more widely available. The potential to replace nucleic acid testing for HIV screening has important implications for both HIV clinics globally as well as for donated blood screening. The door is now open to explore the application of Simoa to many other use cases once thought to be only addressable with nucleic acid testing.

Into the Future

Quanterix recently launched its next-generation platform, Simoa 2.0, which addresses many customer requests and input collected during the rapid growth of Simoa's first year and a half on the market, and includes numerous improvements to assay performance, hardware and software refinements, and overall instrument usability. The platform has also proven to be equally remarkable at detecting DNA and RNA at levels similar to PCR, with the ability to measure both proteins and nucleic acids in a single sample. These new capabilities, along with form factors addressing different market segments and research needs, are being actively developed and advanced at Quanterix. The company believes that its Simoa technology can revolutionize today's sick care into true healthcare by detecting and diagnosing indicators of disease far earlier than ever before and allowing people to lead healthier lives.

Quanterix *is an IQT portfolio company that develops ground-breaking tools in high-definition diagnostics. For more, visit www.quanterix.com.*

REFERENCES

1 Rissin, *et al.*, Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations, Nature Biotechnology, 2010, 28, 595–599.

2 Rissin, D. M.; Wilson, D. H; Duffy, D. C. Measurement of Single Protein Molecules Using Digital ELISA, in The Immunoassay Handbook: Theory and applications of ligand binding, ELISA and related techniques, 4th edition, Wild, D. ed, 2013, Elsevier, Oxford, UK.

3 Rissin, *et al.* Multiplexed single molecule immunoassays, Lab Chip, 2013, 13, 2902–2911.

4 Chang, *et al.*, Simple Diffusion-Constrained Immunoassay for p24 Protein with the Sensitivity of Nucleic Acid Amplification for Detecting Acute HIV Infection, J. Virological Methods, 2013, 188, 153–160.

FROM THE PORTFOLIO

The *IQT Quarterly* examines trends and advances in technology. IQT has made a number of investments in innovative technologies, and several companies in the IQT portfolio are garnering attention for their unique solutions.

Atlas Wearables

Atlas Wearables builds devices and software infrastructure to help anyone take control of their workout routine and fitness content. The company's flagship product, Atlas Wristband, is a true activity tracker designed to optimize indoor and outdoor strength training. Atlas has been featured in publications including *TechCrunch, Fast Company,* and *VentureBeat* for its precise and automated exercise tracking. Atlas is based in Austin, Texas and has been an IQT portfolio company since March 2015. **www.atlaswearables.com**

Claremont BioSolutions

Claremont BioSolutions specializes in disposable devices that provide solutions to what is recognized as the "bottleneck" of DNA diagnostics — sample preparation. The company recently announced a partnership with IncellDx, Inc. to couple their enzyme-free tissue homogenization technology with IncellDx's patented reagents. The combination forms a sample preparation system called IncellPrep Liquid Biopsy, to prepare "universal" cell suspensions formed from either fresh or FFPE tissue. Claremont BioSolutions joined the IQT portfolio in September 2012 and is based in Upland, Calif. **www.claremontbio.com**

IntegenX

IntegenX is a leading developer of rapid human DNA identification technology, next-generation sequencing library preparation systems, and DNA/RNA ambient temperature stability and storage products. The company recently announced the RapidHIT ID System, providing forensic DNA profiles with unprecedented speed and ease of use. IntegenX became an IQT portfolio company in July 2006 and is located in Pleasanton, Calif. **www.integenx.com**

Voxel8

Voxel8 exists to disrupt the design and manufacturing of electronic devices by providing new functional materials and 3D printing platforms. The company has been featured in *TechCrunch* coverage on the recent surge in 3D printing investment and innovation, and was highlighted as "one of the best 9 ideas from CES 2015" by *Fast Company. MIT Technology Review Magazine* named Voxel one of its "50 Smartest Companies of 2015." The company is based in Cambridge, Mass., and has been a part of the IQT portfolio since October 2014. **www.voxel8.co**

IQT®
IN·Q·TEL