WEBINAR | APRIL 14, 2021

Welcome

MATTHEW CONNELLY

Speakers

CAITLIN RIVERS

DYLAN GEORGE

JEFFREY SHAMAN

JAMES FLEMING

Moderators

WILMOT G. JAMES

ALEX N. HALLIDAY

Event presented by

COLUMBIA UNIVERSITY | ISERP
INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AND POLICY
IN THE FACULTY OF ARTS AND SCIENCES

THE CENTER FOR PANDEMIC RESEARCH

Event Partners

COLUMBIA | VAGELOS COLLEGE OF PHYSICIANS AND SURGEONS
PROGRAMS IN GLOBAL HEALTH
PROGRAM IN VACCINE EDUCATION

COLUMBIA CLIMATE SCHOOL
THE EARTH INSTITUTE

APS ACADEMY OF POLITICAL SCIENCE

COLUMBIA UNIVERSITY
DATA SCIENCE INSTITUTE

Advance Warning Systems and Forecasting Outbreaks

THE WIDE SCALE health and societal impact of COVID-19 have thrown into stark relief the lack of coordinated advance warning systems for epidemics and pandemics. This seminar featured public health and policy experts discussing the forecasting of infectious disease outbreaks; where we stand now, and what systems will be developed in years to come. The webinar was co-sponsored by the Center for Pandemic Research at the Institute for Social and Economic Research and Policy (ISERP), Columbia University's Program in Vaccine Education at the Vagelos College of Physicians and Surgeons, Columbia University's Earth Institute, The Academy of Political Science, and Columbia University's Data Science Institute. It was the fourth event of The History and Future of Planetary Threats series convened by ISERP.

The views expressed by speakers are their own and not necessarily those of any organization with which they are affiliated.

MATTHEW CONNELLY[1]: Thank you everyone for joining us. My name is Matthew Connelly. I am the co-director, along with Tom DiPrete, of the Institute for Social and Economic Research and Policy. We are the social science institute of Columbia, and our role is to support and promote research and public engagement. I am delighted to see how our senior research scholar, Wilmot James, has mounted this incredibly impressive series of events on the history and future of planetary threats. He has basically done all the work and I will not try to take any of the credit. If I had any influence, it was in making sure that we try to look at these things not just in terms of anticipating the future, much less confining ourselves to the present crisis, but by taking the longer view—that is, the history of efforts to predict and prepare for different kinds of catastrophic threats.

[1] **MATTHEW CONNELLY** is Co-Director of the Institute for Social and Economic Research and Policy and a Professor of History at Columbia University.

This is the fourth event of the series that we have had this year. We have heard from Ernest Moniz of the Nuclear Threat Initiative. We heard from Fareed Zakaria on ten lessons for a post pandemic world—definitely taking a longer view. Then we also had a symposium on crisis communications and vaccine uptake in fragile African settings. I am delighted to turn things over to Wilmot to introduce this panel on advanced warning systems and forecasting outbreaks.

WILMOT JAMES[2]: Thank you very much, Matt, and welcome to everybody joining us today. Today's meeting will focus on advanced warning systems and forecasting outbreaks, with an overview of the National Center for Epidemic Forecasting and Outbreak Analytics. We are very happy that we have among the best experts in the field with us here today.

I look forward to co-moderating the session with Alex Halliday. Alex is the director of Columbia's Earth Institute. He joined the Earth Institute in 2018 after spending more than a decade at the University of Oxford. During his time at Oxford he was a Dean of Science and Engineering. Thank you, Alex, for co-moderating this session with me.

I would like to introduce the two speakers who will take us through the plan for the National Center for Epidemic Forecasting and Outbreak Analytics. For some background, on the first full day in office, the Biden-Harris Administration announced the intention to create a National Center for Epidemic Forecasting and Outbreak Analytics. The administration also requested funding in the American Rescue Plan to support this effort.

Modeled after the National Weather Service, this epidemic forecasting center would help public health officials anticipate and respond to outbreaks before they grow into epidemics or pandemics. This capability must be developed, it was noted, to ensure that the United States is never caught unprepared again. So, this is a big subject, and we are very pleased to have two of the people involved in the design of this plan.

The first is Caitlin Rivers, who is a Senior Scholar at Johns Hopkins Center for Health Security and an Assistant Professor in the Department of Environmental Health and Engineering. Her research focuses on improving public health preparedness and response, particularly by improving capabilities for "outbreak science" and infectious disease modeling to support public health decision-making. Caitlin participated as an author or contributor in influential reports that are guiding the United States' response to COVID-19 today.

I will also introduce Dylan George, Vice President at In-Q-Tel (IQT). In that capacity Dylan supports technical evaluations of life science and healthcare deals, and drives strategic science and technical vision to strengthen capacity within the United States to counter biological threats from infectious disease epidemics. He served on the Biden transition team supporting national security and foreign policy, and on the agency review team for the Department of Health and Human Services. So thank you, both Dylan and Caitlin, and we certainly look forward to what you are going to discuss with us.

CAITLIN RIVERS[3]: Thanks again so much for having us. We are really excited to be here and share some of the thinking that we have around why this is such an important capability. I want to first start out by laying out a few central challenges that I think epidemiology as a field is facing. The first is that outbreaks keep happening way more than people realize. COVID is not the

[2] **WILMOT JAMES** is a Senior Research Scholar at the Institute for Social and Economic Research and Policy (ISERP), College of Arts and Sciences, Columbia University.

[3] **CAITLIN RIVERS** is a Senior Scholar at the Johns Hopkins Center for Health Security and an Assistant Professor in the Department of Environmental Health and Engineering at the Johns Hopkins Bloomberg School of Public Health. As of August 2021, she is serving as Associate Director on the new leadership team charged to establish the Center for Forecasting and Outbreak Analytics.

only one-in-a-hundred-year event that we have faced. It is about every two years that we face serious public health threats—SARS in 2003, H1N1 in 2009, and more recently, COVID, Ebola, and Zika. These events really do keep happening. And they are not just a grave threat to human life and health, but they also disrupt economic activity. I think that is why we need to change our thinking from this being a once in a while type of situation, to something that can really be a national security threat.

It is not just the outbreaks that make the headlines that we worry about. There are a lot of health security threats that do not even rise to the attention of entering the public consciousness. In 2020, there were 74 disease outbreak news event reports published by the World Health Organization (WHO), and in 2019 there were almost 120. In order to have a disease outbreak news alert issued, an outbreak must be of unknown cause, of significant health concern, of a known cause with a demonstrated serious public health impact, or one of high public concern. So not only are these events common, but they are also quite serious and difficult to control. The COVID pandemic has been ongoing for 16 months, and we have a lot ahead of us. In 2014, there was a serious outbreak of Ebola in West Africa. That outbreak infected nearly 30,000 people and took two years to control. Even now, there is another outbreak of Ebola ongoing in Guinea, and several recent outbreaks in the Democratic Republic of Congo have been challenging to control.

These major events happen often, cause serious extended disruption, and are often very difficult to get back into hand. So, what can we do about them? Medical countermeasures are what most people turn to first when evaluating options to address these health security threats. I do think the development of vaccines is one of humanity's greatest achievements—certainly one of the greatest public health achievements. And so it is not to say that vaccines do not play an important role in therapeutics, but they cannot solve every problem. Research and development is slow and expensive, and that is not the only hard part. Even when we have a vaccine in hand, we find new failure modes over and over again for how we can bring that to bear on emerging infectious diseases. I think medical countermeasures are a tool in our toolkit, but they cannot be the only tool in the toolkit.

Here are a few examples, which we have had to contend with, of a period between when an outbreak emerged and when we had a vaccine to bring to bear. In 2017, there was a major shortage of the yellow fever vaccine that coincided with a major outbreak. WHO had to institute fractional dosing to extend coverage of the vaccine to as many people who needed it. With COVID, even though we have developed a vaccine in record time and well beyond what seemed possible even a couple of years ago, we still had almost an entire year of the pandemic to contend with without vaccines.

I do want to spend one more moment thinking about the specific example of the Ebola vaccine. Thanks to a long history of investing in Ebola research—though in retrospect, not enough investment—there were several candidate products in research and development available to move into clinical trials when the West Africa Ebola outbreak hit. But even with the head start of having products in the R&D pipeline, it was not until the outbreak had largely been brought under control with traditional public health measures that the vaccines were ready to field for clinical trials. Even now, five years later, when we have vaccines that are authorized for use and that are known to be safe and effective against Ebola, we still face challenges in controlling Ebola outbreaks. A recent outbreak in the Democratic Republic of Congo took almost two years to control. Again, that was after the vaccine had completed development.

All of this is to say that medical countermeasures cannot be the only thing that we rely upon. We need to develop and incorporate other capabilities into outbreak response. I think this means we need more innovation. The field of epidemiology is built on data and data analytics, and I think that is where a lot of new innovation can come from even still.

There is an old story about the founding of epidemiology that I think every epidemiologist knows. In 1854, there was a cholera outbreak in London. At the time it was not recognized that cholera is spread through water. It was thought it was caused by bad air. But a man named John Snow had his doubts about the bad air hypothesis. He had an inkling that water had something to do with it. He went door to door, and asked people where they fetched their water from and whether their household had a case of cholera. He put that on a map and he identified clustering around a specific pump. They removed the pump handle, and the outbreak was stopped.

It could be that that is allegorical. But I do think it says a lot about how we can use data and epidemiology to make decisions. I think that is one of the tools that we need to keep in our toolkit, along with medical countermeasure development.

This is where outbreak data and modeling come in. There is so much that we can do with data and modeling to support outbreak response. The most common output is a forecast or the expected future number of cases. A related analysis is nowcasting, which is trying to understand how many current cases there are that have not yet made it through the surveillance pipeline. One kind of forecast that you are probably very familiar with is a weather forecast, which tells us what we can expect in the next few hours, days, or sometimes weeks. We use this information to make all kinds of decisions about whether we should go to the park this weekend or whether we should bring an umbrella. We use that information to change our behaviors and to make decisions in our lives.

But that personal use of weather forecasts is actually a late addition to the weather enterprise. A lot of the original impetus was for economic reasons. Cargo ships wanted to know whether any storms would be met. The airline industry wanted to know whether it was safe to deploy aircrafts based on weather. A lot of the uses of weather forecasting were not foreseen until we already had that capability at hand. I think that says something too about what we can expect for disease forecasting. We can think of a lot of ways that it will be useful now, but I think when we have accurate and reliable disease forecasts, we will find many more uses for how that can be implemented in our everyday lives.

There are various other uses of modeling and data analytics. It is not just forecasting, although I think that is the most evocative. There are other important analyses that we use all the time in disease modeling. Comparing interventions is a common one. Often when you are responding to an outbreak, you do not have unlimited resources. Knowing which interventions are likely to be most effective is valuable for prioritizing, helping decisionmakers, and directing resources efficiently and effectively. I will note here that experienced epidemiologists and public health decisionmakers already have a very good intuition about how to respond to an outbreak—whether it is better to do A or B or C. Even though there are experts who do this intuitively, I think it is still really important to have models, to be able to interrogate your understanding, and to play out different scenarios. It is important to structure thinking around what the questions on the table are, what the possible options are, and how they might stack against each other.

There is one other observation from the 2014 Ebola outbreak that I think illustrates how it is not just fancy data analytics that has value. Improving data quality and availability is also important for making the rest of these fancy approaches work. In affected communities in 2014, the infrastructure for data collection was not well developed, and there were issues with quality. Data elements were missing. They were duplicated. This happens all the time in public health data. It was not specific to that outbreak.

One problem we did not have in the 2014 Ebola outbreak was that all three affected countries collected the exact same data. In advance of that outbreak, there was a case report form developed by the Africa Centres for Disease Control and Prevention in collaboration, I believe, with the United States Centers for Disease Control and Prevention (CDC) that enumerated the different fields that should be collected during the Ebola outbreak—including demographics, exposure information, and clinical information. That little bit of foresight—of putting together this form and making sure that the public health departments had it—meant that the form could be pulled off the shelf when the Ebola outbreak hit. This little intervention really changed the quality and availability of the data that we had from that outbreak. It goes to show that some structure and motivation around putting these puzzle pieces in place can really get us to a better place when a crisis hits. I think that is an important example.

Right now, data modeling and analysis work in the United States consists of a robust community of academics who volunteer their expertise when there is a crisis. It is really that wealth of expertise, and that generosity to step into the breach, that has powered the COVID response. It also powered the Zika response and the Ebola response. Although I am very grateful for that expertise, I do not think we can carry on like that. That is not how we structure national defense. It is not how we structure weather forecasting. A capability that we rely on so heavily to guide our national interests is something that we should institutionalize. We should make sure that it is a resource and fit for purpose.

That is why I am so excited about this National Center for Epidemic Forecasting and Outbreak Analytics. As Wilmot mentioned, it was announced in National Security Directive 1 that the Biden Administration intends to establish this center. $500 million was appropriated in the American Rescue Plan to the CDC in order to implement this center. Some of that money is also for public health data monetization, which goes back to my Ebola example about how important it is to invest not just in the fancy analytics, but also in the data.

I am really excited to see the center come into existence. I am excited to see how it changes and supports our outbreak response, not just for COVID, but also for the next crisis and the one after that. This is going to be an important permanent capability for our long-term interests.

After every crisis, not just in public health, but generally, there is a period of reflection where we assess what we can do better next time. That is because we know that there will be a next time. There always is, and we can expect another infectious disease threat in this case as well. It is excruciatingly clear that it is in our national interest to be ready. I think that outbreak forecasting and analytics is key to that readiness, and I think this center is going to help us get there. I will pass it over now to my colleague, Dylan George.

DYLAN GEORGE[4]: Thank you very much, Caitlin. Thank you, Wilmot, for putting this together and for the opportunity to be with you all today. I just wanted to add a few framing comments to what Caitlin has said.

We live in an age of pandemics. Pandemics are, and will continue to be, a major national security threat. This administration will face more, if not multiple, public health emergencies during its time in office. It is interesting that over the past 10 to 15 years, we have had to try to convince people that we live in this age of pandemics. I think it is painfully clear now that this is true. Given the constant threat and a significant potential impact, we need to meet this current moment and we need to prepare for persistent future threats going forward.

We are at a bit of an inflection point in public health, I believe. In recent decades, governments, particularly the United States government, drove objectives and funding for global health goals. COVID-19 has shown that single point failures are costly, and an ecosystem of stakeholders is needed. This definitely includes governments, but it also includes non-governmental organizations, the civil society, volunteers, and the private sector.

[4] **DYLAN GEORGE** was Vice President at In-Q-Tel (IQT). As of August 2021, he is serving as Director for Operations on the new leadership team charged to establish the Center for Forecasting and Outbreak Analytics.

Caitlin and I had the honor of helping in the Biden transition with preparing for COVID-related policies. In that time, we had the privilege of meeting with many companies, organizations, and volunteers. I was particularly humbled by how many of these people were working, and how many of these companies were investing at risk, or even against their own interests, to help with the pandemic response. It was absolutely inspirational. It was very clear to me in those discussions that significant capacity and capabilities lie in the private sector and civil society. And we need to find ways of taking advantage of that to more effect going forward.

Seeing those people meet the moment helped me realize that reinforcing this ecosystem will help us with future pandemics that are assuredly coming down the way. Incidentally, in the United States, we have explicitly accepted the need for private sector involvement in manufacturing vaccines and drugs. We turn to the private sector—Johnson & Johnson, AstraZeneca, Pfizer, Merck—to manufacture vaccines at scale. Without a doubt, the government has a clear role in driving the funding, the requirements, as well as in de-risking the science. But the private sector is manufacturing the products. So there is this really interesting public–private interface. To manufacture at scale, we needed the skills and the capabilities that were resident within the private sector.

COVID has demonstrated that we similarly need to be explicit about the need for private sector involvement in data systems to guide public health responses. If vaccines are mission critical, and we lean on the private sector to build them, why are we not doing the same for mission critical capabilities like data systems that guide public health interventions? This is going to be a pressing question in the months and years ahead.

One suggestion, that I would like to at least posit, is to learn from what enabled making vaccines for pandemics. After 11 September 2001, the United States government realized that we needed to make vaccines faster, especially those that fell prey to market failures. This was mainly the pandemic vaccines, like what we are seeing with COVID. This would require the federal government to work with the private sector in a different way than was done before. The Biomedical Advanced Research and Development Authority (BARDA) was created in part to enable these public-private partnerships that would be needed to push forward these critical vaccines. I used to work at BARDA, so I have a soft spot in my heart for the organization and their mission. I would be the first one to admit that there are warts associated with BARDA. It is not a perfect organization by any means, but it is very good at engaging the private sector and moving things forward in a public way for a public good.

The federal government has struggled with data technologies and attracting data scientists. This is a well-known problem. It is no different for data systems within public health. These systems are legacy systems. They are decades old. These systems need to be modernized, and the private sector has to play a major role. I have been enamored with this idea of finding new ways for the federal government, in particular, to engage the private sector on data technology. It is not going to be easy. It is not going to be straight forward. But I think that we need to meet this moment and to move forward, such that the private sector and public health can work together more effectively. So perhaps we do need a BARDA for data technology and public health.

We live in this age of pandemics. We also live in an age of emerging technologies. As Caitlin pointed out—and we are all painfully aware from our common experience with COVID—when a novel pathogen emerges, we will not have drugs or vaccines to protect us. As great an accomplishment as it has been to make the vaccines at scale, and as fast as has been done, we are still over a year into the pandemic and people in communities are fatigued and hurting. Vaccines and drugs, but only recently, have been helpful and well received in this fight. The things that have been helping us fight this pandemic are exactly what Caitlin talked about—public health interventions such as wearing masks, washing hands, watching our distance, and staying at home.

These measures have kept us safe. They are critical tools in the fight. Data and analytics can allow us to guide these interventions for maximal effect and hopefully minimize the adverse effects that we see, particularly economically.

Unfortunately, the current data systems have slowed our response and they are challenged in trying to keep pace. The country has stumbled in using data and analytics to guide our response. It is humbling that although we live in a country that has transformed the world with data technology, our healthcare and public health data systems remain decades old and not fit for purpose. I have been inspired and humbled at the volunteer efforts by multiple people and groups. Two exceptional efforts have been led by the Johns Hopkins Coronavirus Resource Center and The COVID Tracking Project. We were all using their sites to track the pandemic for a really long time, and we owe them a debt of gratitude for their work and their sacrifice. They sacrificed to do that for us. Their efforts helped us to understand what was happening with the virus. So, we need to capitalize on efforts like those and take advantage of the technologies that we can find.

As Wilmot mentioned, I am very encouraged that the Biden Administration and Congress have supported the idea of strengthening the overall public health data pipeline, and specifically this National Center for Epidemic Forecasting and Outbreak Analytics. This center, we do believe, will help people make decisions faster, better, and stronger, so that we can maximize the impact of responses and minimize the adverse effects. We need to figure out how to do this more effectively, and the center will be a focal point for bringing that together. This is important work. As Caitlin pointed out, currently a lot of this work is being done by volunteers in academia, in civil society, and in the private sector. These folks are typically not paid, and not organized or formally tasked to support these efforts. The center would bring them together, professionalize this service, and make it able to move at speed and scale in a completely different way.

There are a couple of key components that we think about with a National Center for Epidemic Forecasting and Outbreak Analytics. We have chatted with colleagues that have worked in weather and hurricane forecasting, and there are a couple of key components that have allowed them to be successful. For weather forecasting, they have developed systems that operate at three levels: improving the science and developing new capabilities, translating that science into technology, and then, lastly, interpreting the results at the local level. Organizations like the National Center for Atmospheric Research focus on improving the science of forecasting. We have translated the best of science into operationalized technology in the National Weather Service to forecast daily. Then we have a system of local meteorologists that are interpreting the models for local consumption and decision-making. This system, albeit simplistically described, allows for improving the science, the technology, and the communication of results, so people can use that information to protect themselves and their loved ones. For the National Center for Epidemic Forecasting and Outbreak Analytics to be successful, it needs to be focused on helping decisionmakers make decisions bigger, faster, and stronger. If it does not accomplish that function, then we need to reallocate resources to something that can. So, this center should live or die on its ability to do that.

There are some competing ideas about analyzing and anticipating emerging events, versus analyzing ongoing outbreaks. Many of us know that the most infectious diseases in humans come from pathogens that circulate in animal populations. Ebola, Zika, and many coronaviruses originally came from animal reservoirs that spilled over into humans. Researchers are working to understand the spillover dynamics and eventually anticipate spillovers—particularly the spillovers that would cause major outbreaks. This incorporates ecological, social, behavioral, virological, and immunological components, among others, to try to understand those major drivers. If we can identify those spillover dynamics, then we potentially could stop the spillover from happening. It is an area of research that many people are focusing on.

When I was in the White House during the Obama Administration, we called efforts to anticipate spillover and disease emergence as "prediction." It is just a term of art that we used, more than anything else. This kind of work was distinct, at least in my mind, from using analytics to understand an ongoing outbreak like how Caitlin described. The projections of an ongoing outbreak, and looking around the corner as to what would happen, is what we called "forecasting." The full spectrum of alerting, risk characterization, scenario planning, and forecasting is what we call "outbreak science." I think that we will make the most progress the fastest on forecasting and in supporting decision-making for outbreak response, rather than on prediction of spillover events. For me, the near-term focus should be on forecasting and outbreaks.

Having said all of that, I am exceptionally encouraged that the Biden-Harris Administration has mandated efforts and included resources for funding the Data Modernization Initiative and the National Center for Epidemic Forecasting and Outbreak Analytics, as well as for more genomic sequencing, so that we can understand what is happening with the variants. I am also very encouraged that, as the administration put out last week, they are recommending an increase in funding generally for the CDC in their most recent budget proposal. These initiatives can transform data and analytics—our ability to effectively guide interventions. These capabilities will put us in a much better position to respond to future pandemics that we know will come. With that I will conclude. I very much look forward to learning from Jeff and Jim's comments, and then chatting with you as we go into the Q&A. Thank you very much.

JAMES: Thank you very much, Dylan and Caitlin, for walking us through the fundamental importance of improving data collection and refining our ability to be more predictive and more anticipatory when it comes to disease outbreaks. That was really fabulous. Dylan and Caitlin are also quite modest—they played leadership roles within the U.S. policymaking environment on a number of occasions. If you have not yet read the Biden-Harris plan for dealing with COVID-19, look at that last paragraph when you do. They had a lot to do with the design. So, thank you both very much for your leadership and for your expertise, and for your contribution on the academic and intellectual front in dealing with two key issues.

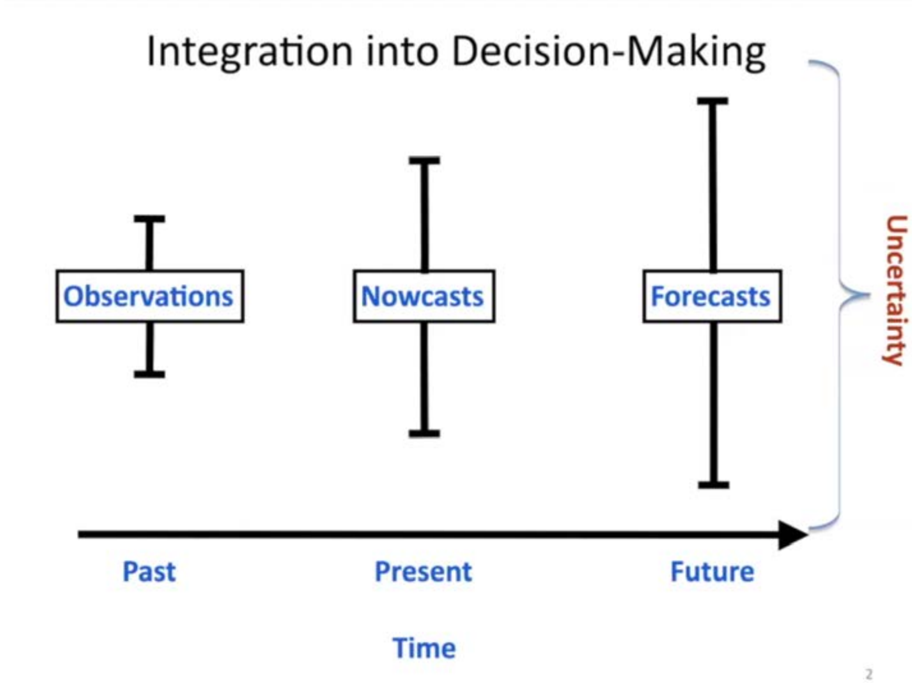
The United States is playing catch-up. There are other countries which have similar institutions—Europe, Hong Kong, and a few other places. The reason why the United States is catching up is because of the vision of the Biden-Harris Administration. This is a really nice moment for us to be having this conversation.

I am now going to hand it over to my colleague, Alex Halliday, the director of Columbia University's Earth Institute.

ALEX HALLIDAY[5]: Thank you very much, Wilmot. This is a great and very interesting panel to be associated with. It is my great pleasure now to introduce Jeff Shaman, who is a professor in Environmental Health Services. He is also director of the Climate and Health program in the Mailman School of Public Health at Columbia University. As of this year, he also is the new chair of the Earth Institute faculty. Dr. Shaman's background is in biology, but he worked on modeling climate variability and its impacts for his Ph.D. Today he studies the environmental determinants of health with a strong focus on infectious disease, including COVID-19. He is going to talk to us about improving infectious disease forecasting.

[5] **ALEX HALLIDAY** is the Director of Columbia University's Earth Institute and Founding Dean of the Columbia Climate School.

FIGURE 1



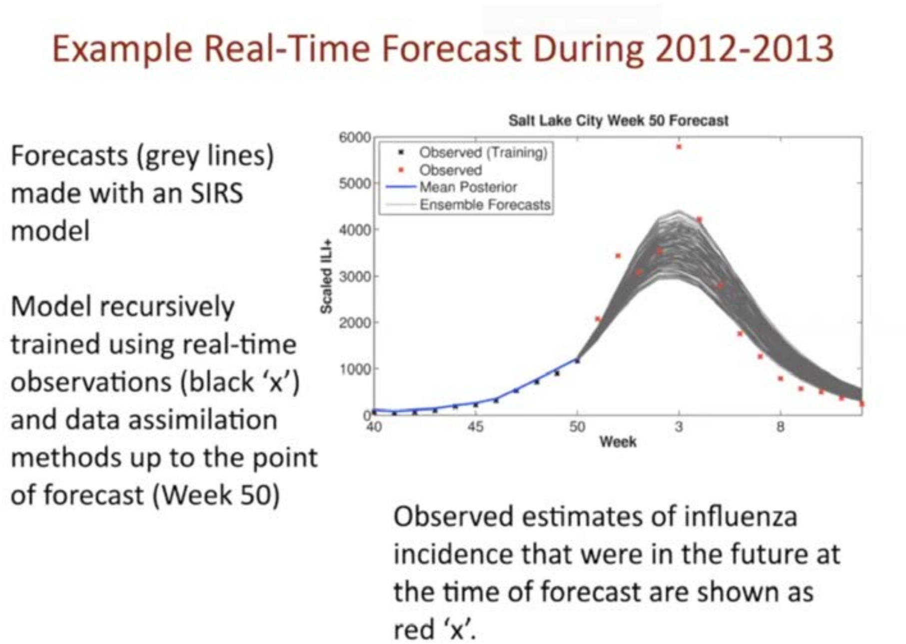
JEFFREY SHAMAN[6]: Thank you, Alex, and thank you, Wilmot and Matt, for organizing this. Dylan, Caitlin, and Jim it is a pleasure being here with you. I was asked to speak about improving infectious disease forecasting in the context of forecasting that Dylan described.

As Caitlin alluded, I imagine much of the center will be organized around data—data acquisition, organization, sharing, as well as analytics—and will leverage a variety of statistical machine learning and mathematical approaches to understanding those data. A lot of the work that goes into understanding infectious disease leverages mathematical methods and mathematical modeling. There are three general problems that we try to interrogate with these models of infectious disease. One is to simulate and infer or estimate the properties that guide the transmission dynamics of these diseases at scale. The second has to do with using those models, be they mathematical, statistical, or otherwise, to project future outcomes—such as how much disease burden there will be, how much transmission there will be, and where a pathogen will spread. So that falls in the scope of forecasting. The third is to explore counterfactuals, interventions, and the ways in which we might intervene or blunt the effects of an infectious disease, as we anticipate its future impact.

I am going to focus on the second of those, which is forecasting. Forecasts lie at one end of the continuum of information we have regarding infectious diseases (Figure 1). That information starts with observations, which provide indications of what has happened in the past. Incidence of disease, cases, deaths, and transmission hotspots, for example, are recorded in an observational record. The continuum progresses to the present, where the information becomes much more gappy. This is because there are many lags in the reporting of information. Very often we use statistical methods to try to supplement the limited observations that we have in real time. We call those nowcasts. It is an effort to try to fill in the gaps and create a more complete picture of what is going on in the here and now. The continuum then extends into the future with models—that we use dynamically, statistically, or otherwise—to project what conditions will be like in the future.

[6] **JEFFREY SHAMAN** is a Professor in the Department of Environmental Health Sciences and Director of the Climate and Health Program at the Columbia University Mailman School of Public Health.

FIGURE 2

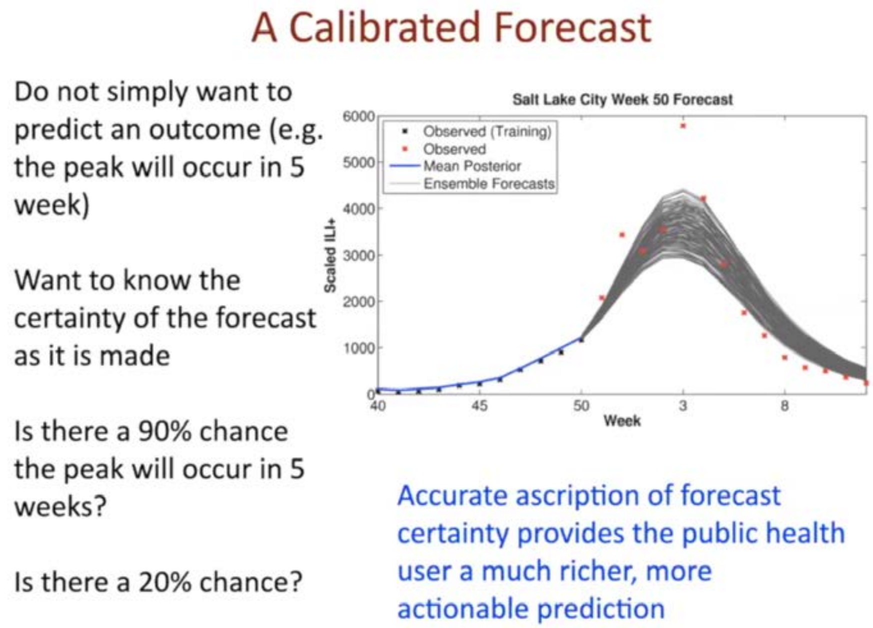


We have error bars drawn around the observations, nowcasts, and forecasts in Figure 1 that tend to get a bit larger as you move from the past, to the present, and into the future. The uncertainty of the estimates—and our understanding of the information about the past, or the observation—tends to be more constrained. When you move to the present, the nowcasts, it tends to increase a bit. The uncertainty tends to be greatest when projecting what is going to happen in the future because there are many factors that we may not understand. That does not always have to be the case. There can be forecasts that are very accurate, and there are observations that are frankly lousy with enormous uncertainty or error around them. It is actually quite variable, but as an overall tendency, uncertainty increases from past to future. Any judicious, intelligent use of these pieces of information wants to account for that uncertainty. You would like to know if your observations are good or highly uncertain, and you would like to draw upon that knowledge when making a decision based on that information. If you know that the observations are not good, you will likely discount them more. If you know they are of high quality, you will pay more attention to them. The same thing goes for the nowcasts and the forecasts.

As part of the scientific inquiry—in our efforts to improve forecasting, nowcasting, and even observations—we want to shrink those error bars. So, it is the job of the scientific community, and hopefully this national center, to start bringing those error bars in more tightly, so that we have more reliable, calibrated, accurate information at our disposal—be it for the past, present, or future—to understand the disease threats that we have experienced, the conditions right now, and the threats we will be confronting in the weeks and months in the future.

I am going to give you an example of what a forecast looks like (Figure 2). This is an example of an operational real-time forecast for Salt Lake City that my group generated for seasonal influenza during the 2012–2013 season. It is week 50 of 2012, around 15 December. The forecast was initiated roughly on that date. The black X's are observations of influenza-like illness plus, which is an estimate of influenza in the community, shown for Salt Lake City from weeks 40 to 50. The red X's show observations that we did not have in hand at the time that ultimately, as this is in the past, we now know. The blue line is a posterior fit by an ensemble of models, or multiple simulations, each one of which is itself an ensemble.

FIGURE 3

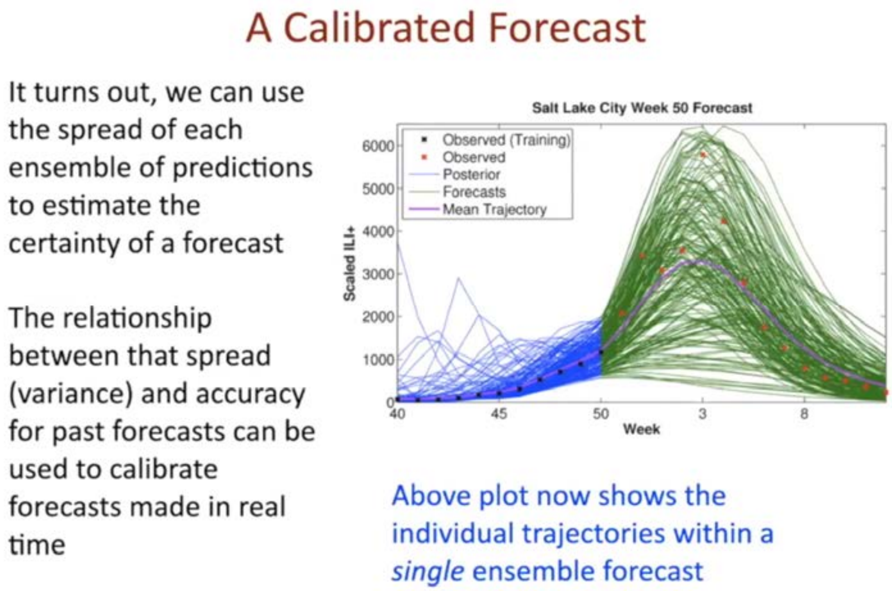


The gray lines are the forecasts—the projections into the future running out about 10 weeks that we made with each of these 200 member ensembles. There are 150 of these 200 member ensembles. Each gray line depicts the mean forecast trajectory. These lines predict that the peak will occur in about five weeks, which is, in fact, what took place. The forecasts depict the overall duration of the outbreak. The area under the curve, the total infectious burden, is also captured by this forecast. Now that peak, that red X, is well above the gray lines, but they are the mean of an ensemble. As I will show you in a moment, they are statistically well captured by this prediction.

We can make this prediction for something that is recurrent like seasonal influenza, for which the properties are stationary and do not vary that much. Here, it is great to be able to predict at week 50 that there will be a peak of influenza in Salt Lake City in five weeks, and be correct in that forecast. But we would like to do more than that. We do not simply want to predict that flu peaks in five weeks. We want to know the certainty of that forecast. This is very analogous to weather forecast. The National Weather Service does not simply tell you it will rain tomorrow. They say there is an 80 percent chance of rain tomorrow, or a 20 percent chance of rain tomorrow. We want that kind of calibrated statistical reliability, which we have come to know and trust in weather forecast, from our infectious disease forecasts (Figure 3).

It turns out we can begin to pull that information by decomposing each one of those gray lines. If I pick one of those gray lines and tease it apart to show the underlying full ensemble, it appears a bit more like spaghetti (Figure 4). The ensemble spans the observations in the future, which statistically is very good. But if we look at the disagreement among the individual projections, we can use that information to make an estimation of how certain we are of the forecast. When those lines are more in agreement, there is higher certainty in the forecast. And when there is more divergence in the green lines, there is less certainty in the forecast. We can use that to say there is a 70 percent chance that flu will peak in five weeks, versus a 10 percent chance that flu will peak in five weeks. Both are giving the same mean prediction, but the certainty associated with those two forecasts is very different. One may be actionable—upon which an organization, a hospital, or a public health agency may want to respond. So, we have the capability of generating these types of forecasts.

FIGURE 4

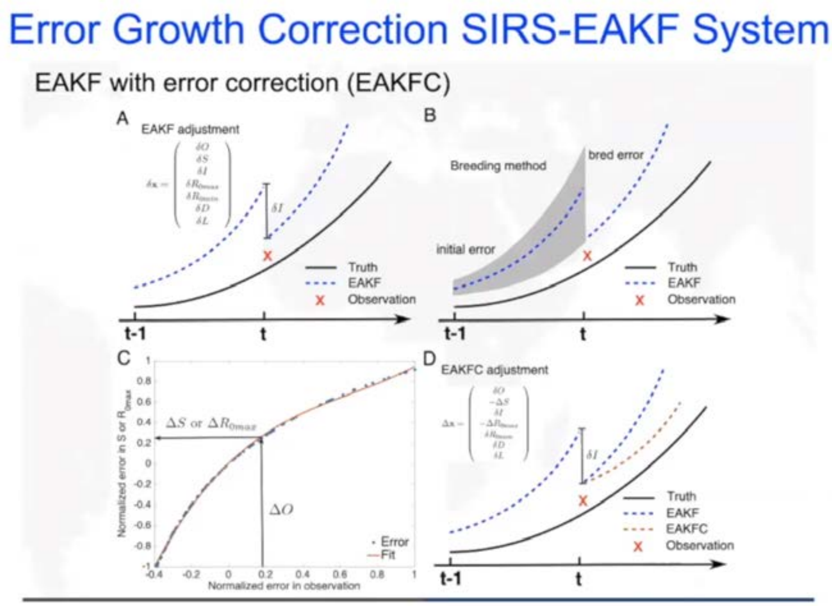


We have certainly seen for COVID-19 what I call projections, rather than forecasts, because there is so much that civil society is doing to try to disrupt the pandemic. It is very hard to predict the policies, actions, and levels of compliance there will be from community to community in combating the disease. With any forecasting or projection system like this, both in academia and certainly through this national center, we want research to advance its capabilities further and to develop more reliable forecasting models. We want to try all sorts of alternate model forms, such as statistical, mathematical, hybrid, age stratified, within host, multiple strains, and spatially explicit or cross scale models. We want to improve the optimization techniques that we use to fit those models to past observations in order to generate a reliable calibrated projection. That includes investment in what are called data simulation methods—sequential Monte Carlo techniques, Bayesian inference, and other means of fitting nonlinear systems to observations.

We want to improve the observations as well. That is a very big issue. We need to access private or novel data streams. We need to get these organized and reduce observational error. We want to improve the availability of these data in real time so that the work of nowcasting is reduced, and so that there is a richer compliment of data available to support forecasting future outcomes and understanding present conditions. Lastly, we certainly want to deliver and archive the forecasts for integrated use and decision making. Both technological and instructional investments will need to be made. We will need to work with public health, medical, and business sectors to integrate the use and adoption of these models and projections in a sensible and informed way that does not oversell or undersell their capabilities.

I will give you a couple of examples of this. One might be a study of nonlinear error growth in these models (Figure 5). The systems we are depicting, much like the weather and the atmosphere, are nonlinear and they have some funny dynamics that need to be understood. We need to understand how error grows within these systems in order to control for, compensate, and handle those system behaviors when making forecasts. This is an area that has been invested in considerably in numerical weather prediction, and we need that kind of investment for infectious diseases as well.

FIGURE 5



Pei and Shaman, 2017

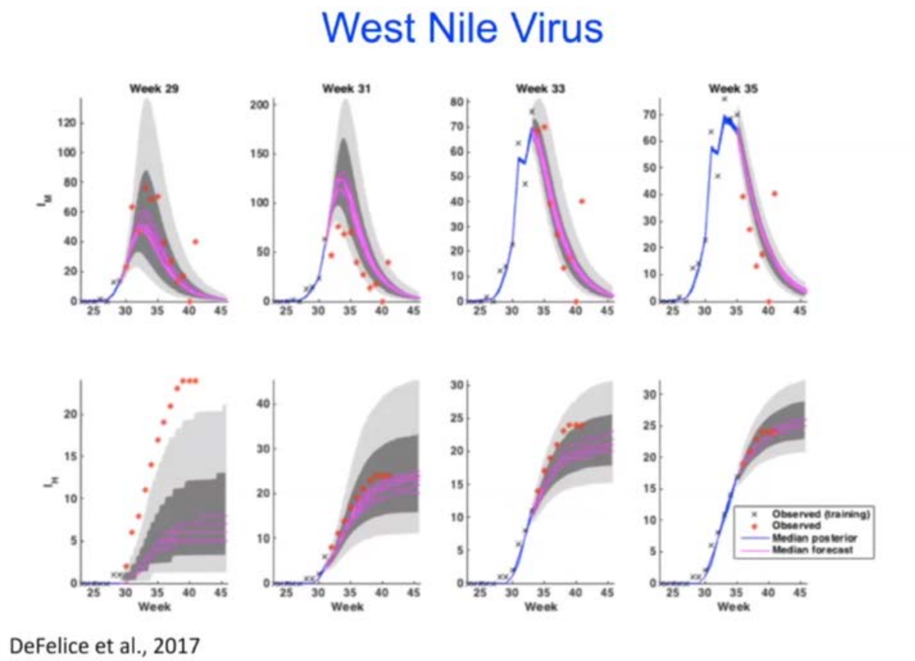
We also have big issues with observations. It has been alluded to by both Caitlin and Dylan, and I will give you an example here from my own work. This was a study that was led by Nick DeFelice of Mount Sinai, who was a postdoctoral scientist with me at the time, to develop a retrospective system for forecasting West Nile virus. The virus is an endemic arbovirus in the United States at this point, and it emerged back in 1999. We were able to show retrospectively that we can make forecasts using mosquito infectious pool data and data of human cases well in advance of the burden of disease. Shown in Figure 6 are observations of infections of humans and forecasts made—the black X's are observations in hand and the red X's are future cumulative cases. You can see the predictions at different points in time as you go from left to right from week 29 through week 35. By the time you get to week 31, which is well in advance of when most of the infections took place and were hospitalized, we are able to predict the West Nile virus burden in this county.

However, we can only make this forecast retrospectively because all the data are in hand. In reality, West Nile virus, even though it is a notifiable disease, has a delay of two to 12 weeks between a person being diagnosed or tested for West Nile virus, and that infection being officially confirmed and made available through public health channels. A three-month delay is much too long to support real-time forecasting. So, we are not capable of making accurate West Nile virus forecasts because the data are not operationally available in real time. We would need to accelerate those pathways for returning information on disease burden.

One of my favorite graphics shows a time series of the history of numerical weather prediction over a 60-year span, starting roughly in the mid-1950s (Figure 7). This graph stops at about 2013, so the plot is a little bit old, but the trend continues. It shows a field that is halfway up the atmosphere over North America and the ability of the NCEP Operational Forecast to predict that field. Skill increases moving upward on the y-axis, and 36-hour and 72-hour forecast skill is shown.

The thing that is important to look at is that monotonically, the skill of these forecasts—their accuracy and ability to deliver useful information—increases over time. It has increased over

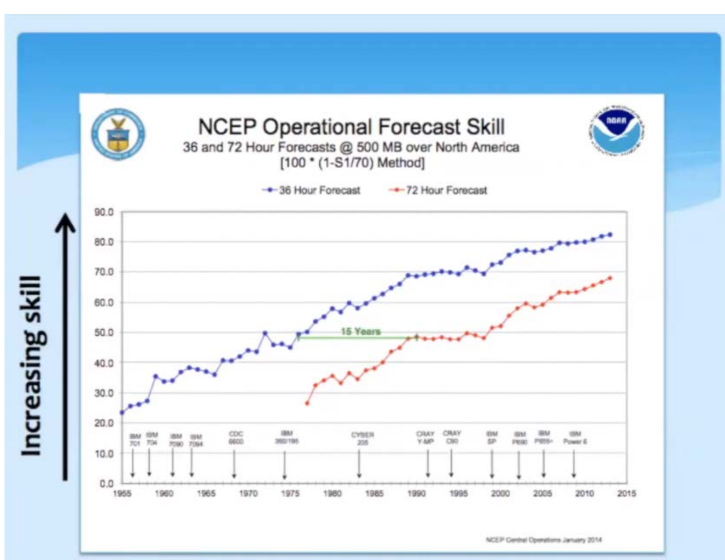
FIGURE 6



time because of investment. We have a National Center for Environmental Prediction that is part of the National Oceanic and Atmospheric Administration, and linked to NCAR and the National Weather Service. They have been investing in understanding the atmosphere and its dynamics at a fundamental level, as well as in integrating academic research, building better models, improving computational capacity, building data simulation methods to ingest observations, and improving the observational network. All those things have happened over this 60-year span of time to greatly increase the skill of weather forecasting. Those forecasts are appreciably better than they were, for instance, when I was growing up.

We need the same kind of investment for infectious disease forecasting and a national center will allow the operational investment, consolidation, archiving, and systematic assessment needed to improve those forecasts, so that we can keep the United States and other parts of the world safe. I will stop there, and I thank you for your attention.

FIGURE 7



HALLIDAY: Jeff, thank you so much. That was fantastic and very inspiring. It was nice that you ended up on the note of history, because that takes us very neatly to our next speaker, who is James Fleming, the Charles A. Dana Professor of Science, Technology, and Society at Colby College in Maine. Jim's background is in astronomy, then he did atmospheric science, but he gained his Ph.D. in history from Princeton. His books include *Meteorology in America, 1800-1870*, *Historical Perspectives on Climate Change*, *The Callendar Effect*, *Fixing the Sky*, *Inventing Atmospheric Science*, and *First Woman: Joanne Simpson and the Tropical Atmosphere*. Jim is a fellow of the American Association for the Advancement of Science and the American Meteorological Society. He is founder and first president of the International Commission on History of Meteorology. His opinions are sought by Congress, National Academy of Sciences, and the International Panel on Climate Change. Jim is going to talk to us about weather forecasting and the historical perspective.

JAMES FLEMING[7]: Thank you very much, Alex. It is really a pleasure to be part of this group because it is so inspirational. I thank the organizers for this opportunity. The history of weather forecasting is just over a hundred years old in the modern sense. In 1904, Norwegian physicist Vilhelm Bjerknes cast the problem of weather forecasting as a problem in mechanics and physics. He said that what we need is a *sufficiently* accurate knowledge of the state of the atmosphere at the initial time, and a *sufficiently* accurate knowledge of the laws according to which one state of the atmosphere develops from the other. I emphasize the term "sufficiently" because perfectly exact measurements are not possible, and "sufficiently" accurate forecasts are the best that can be expected in a complex, non-linear system.

Bjerknes called knowledge of the state of the atmosphere "diagnosis." The development of the future weather he called "prognosis." He was using medical analogies, but was dealing with physical laws. His statement was a statement of faith in rational mechanics, that the equations would give him the next step and state of the atmosphere. He was able to raise certain fundamental questions about how far into the future we can see. The time dependent equations of atmospheric motion, and I think the group here is very aware of that, are an intertwined tangle of non-linear influences. I have called them the Gordian knot of meteorology. If you could untie this knot, you could provide prevision of the weather for 10 days, for seasons, and perhaps for next year. But these equations have no exact solution.

Realizing that hand computations would proceed slower than the weather would be changing, in 1919 the Bergen school developed a graphical method of weather analysis and forecasting. They started with observations, and their quest was to have a uniform system of observations of a sufficient density, lots of them, taken on a regular plan. But then the next step is analysis of the data. The most famous book in the field of weather forecasting is *Weather Analysis and Forecasting*, because the precision analysis of the observations is really important. Then you need a model to impose upon your analysis. They had the air mass analysis model and the polar front cyclone model in Norway. All three elements—observation, analysis, and modeling—are necessary in both the theory and practice of forecasting.

Responding to Dylan's comment about societal needs, weather forecasting was co-produced with societal needs. That is a very important aspect of this. In the eighteenth and nineteenth centuries the needs of agriculture came first, and then shipping and commerce. Then in the twentieth century aviation and space systems were increasingly driven by military and security needs. Large corporations involved in the national weather enterprise include Lockheed Martin, Ball Aerospace, and others that developed the information and space systems.

[7] **JAMES FLEMING** is the Charles A. Dana Professor of Science, Technology, and Society at Colby College.

I wanted to share a couple of reasons for caution, as I read the original article in *Foreign Affairs*, because the Gordian knot has no analytical solution and only approximate graphical solutions. It is based on well understood physical quantities, like pressure, temperature, moisture, and the like, but it is not based on rather poorly understood human transmission factors, as we are dealing with in the case of pandemics.

In the middle of the twentieth century, some transformative technologies, like Alexander's sword, began cutting into the Gordian knot of prevision. Rather than untangling it, they cut through it with general use technologies, which were not created or built by the meteorologist. Radar and satellites provided new observational tools, and then digital computer models crunched the numbers many times faster than the weather changes. You can do the computations at light speed.

Jeff showed the 500 millibar computer forecast and this is where the models are best, because that is where the flow of the atmosphere can be assumed to be friction-free. There is no heating to take into account, no topography, and no moisture to amount to much of anything. So, it is a much simpler computation than trying to predict the weather at the surface. This is upper air forecasting that was done from the 1950s. Even though we were making progress—and we are getting better at combining both the observations and the models to run in a sequence iterated just about every microsecond—there is a new Gordian knot of meteorology emerging.

It emerged in 1960, involving a computational sensitivity to initial conditions. This was what Ed Lorenz brought us with chaos theory, which holds that small differences in initial conditions can produce very great differences in the final outcome. Even today—with the progress we have made, and we have made quite a bit—there is currently a seven-to-10-day theoretical limit on predictability of the weather. The weather forecast might be quite accurate for a day or two, and not too bad for several more days, but the accuracy takes a nosedive after that.

I also wanted to share reasons for optimism. It is not entirely a cautionary tale. Bjerknes was undaunted at the time that he could not prepare a forecast because he was more interested in proof of concept than in practical applications. The practical results came later. Speaking in 1913, he noted, "It may require many years to bore a tunnel through a mountain. Many a laborer may not live to see the cut finished. Nevertheless, this will not prevent later comers from riding through the tunnel at express train speed." Tangible results soon followed with the Bergen school methods, allowing for useful forecasts, at least for several days. Now modern technological methods allow us ready access to real-time graphical nowcasts on our computers and smartphones.

Back to a cautionary note to finish up here. Weather systems are exceedingly complex, as I mentioned, but they involve measurable and well-defined quantities, subject to physical laws. Pandemics are orders of magnitude more complex. They involve mutating and poorly understood viral agents, subject to both biological and physical principles, as well as the vagaries of human behavior, both individual and collective. Perhaps we stand metaphorically, as Bjerknes described, with pickaxe and shovel, ready to tunnel through a mountain. Perhaps in relatively short order, we will have a surveillance system in place for pandemics, as we have for tropical storms. Yet, we still have this cone of uncertainty. For example, a modern category three storm like Hurricane Katrina became a heavy rain event within 24 hours. Yet, it still devastated low lying areas of New Orleans, in large part due to the failures of the human-built infrastructure. Perhaps global climate modeling and prediction provides an apt analogy since it, like pandemic modeling and prediction, is a wicked problem with diverse unbounded social entanglements.

The national weather service was started in 1870, and it was a military outfit. NOAA, our modern National Oceanic and Atmospheric Administration, dates to a reorganization of federal services in 1970, a century later. But the federal agency for weather prediction was preceded and supported by decades of atmospheric research. This was mentioned too by the other speakers. A national center for forecasting and analytics, which is really a wonderful idea, also needs a national center for research, perhaps embedded in it or perhaps associated with it. In atmospheric science, the National Center for Atmospheric Research (NCAR) is supported and governed by a massive University Corporation for Atmospheric Research (UCAR) involving over one hundred different coordinated university research groups. By analogy, we need to learn how to understand and monitor pandemics and not just aim for prediction and control. Although, we need to have that as well.

Seeing Jeffrey's opening graphic reminded me that there are now some opportunities for atmospheric scientists to get involved in the pandemics. A group I am aware of is working on studying droplets. They took their cloud physics instruments and sneezed into them to see what they could do to measure the behavior of sneezes. They measured a five-micron droplet with their cloud physics equipment. Cloud droplets are typically much larger. They are 20-microns in size. But clouds are able to float sometimes for very long distances across political and social boundaries. And clouds rain and snow on everyone. So, the microphysics of aerosol suspensions is an important area for cross-over research. I think there is a lot of hope here, and a lot of initiative, to move forward in predicting pandemics, but there is a lot more behind the history of weather and climate prediction by way of analogy that we need to put into this equation. Thanks for your time, and I will turn it back over to the group.

HALLIDAY: Thanks very much, Jim. Now, we will engage in the Q&A. Wilmot has been collating some of the questions that have been coming in, and we will be asking the panelists for their responses.

JAMES: This first question is directed at Dylan and Caitlin: How do you envision the National Center for Epidemic Forecasting and Outbreak Analytics will coordinate with the relevant state and city institutes—such as the proposed $20 million New York City Pandemic Response Institute—when it comes to outbreak modeling? How would it be structured to leverage expertise in academia and the nonprofit sector?

GEORGE: I will start off and then Caitlin can join in. First off, Caitlin and I are not in the federal government. This is just what we think it should be, just to be clear on that. As I mentioned in my comments, for public health decisions, the rubber hits the road at the local level, at least in the United States context. We need to figure out how to support mayors and governors in making decisions better, faster, and stronger in some way. There has to be a very strong interface at the state and local level with the results that are going forward and efforts to interpret those results. That is why we talked about the R&D component, the operational component, and the decision support component. That is the measure of success. Are we doing that? Are we helping those decisions move forward more effectively? It is not clear how that is going to materialize, but that is the intent going forward.

Now to the question about how academia, non-profits, and civil society are going to interface with this. There is the intent to engage both academia and civil society in improving the science, in developing out new capabilities, and in testing new capabilities in the localities. In framing the overall structure of this, there is a lot of a desire to figure out that process and how to bring all the stakeholders along to define this. And again, I am speaking as an individual, not as somebody that is representing the federal government in any capacity. I will pass it over to Caitlin if there is anything else that she wanted to add.

RIVERS: I would just add that, right now, a lot of these activities are being run out of academia. I think our vision is to keep doing that, because that is where a lot of the expertise is, but let us make sure that we also have the operational capability to pivot to response activities. Researchers are good at doing research, but there is a lot of space left for people who are great at interfacing with state and local public health officials, and who are great at helping to support decision makers. It is really about growing and building that academic base into something that is able to be put to a response mode more easily.

JAMES: Thank you very much. There is a follow up question that read: I am interested from a venture capital investor perspective to invest in companies that are working on solutions for advanced warning systems and forecasting outbreaks. Will we see the data to help investors decide what companies or researchers to back? This question is not directed to any one of our individual speakers, but for anybody who would like to take that up.

GEORGE: The organization that I work at right now is In-Q-Tel, which is a strategic investor. We operate very similarly to a venture capital firm, but we invest in venture-backed startup companies for the federal government. Instead of getting return on investment, we are focused on enabling technology transfer to the federal government in some capacity for a range of different technologies. I do think that there is an ecosystem of data systems and data companies out there that are going to be critical going forward in improving the whole data supply chain. How do we collect data more effectively? How do we clean and aggregate data? How do we then analyze it? How do we visualize it and then present it? How do we interpret it more effectively for decision support?

All along that data supply chain, there are going to be very useful component parts and technologies. I think the biggest challenge and frustration is going to be how those technologies and companies can interface with the healthcare system and the public health system within the United States. That has always been a bit of a challenge because it is not a full market system. It is a warped system that the payers and providers are dis-aggregated from one another in moving forward.

I am very encouraged based on what I have seen at In-Q-Tel, and based on some of the trends that we have seen over COVID. There will be more and more opportunities to move forward. I think that the interest in telehealth, at-home sampling and testing, wearables, and digital technologies for improving clinical trials, or just even monitoring individuals, are all showing that there is going to be lots of opportunity for venture backed startups for technologies moving forward.

HALLIDAY: I would like to ask a question on the social sciences side: How will the work coming out of the center coordinate and/or align with behavioral sciences to understand communities' propensity to respond and, therefore, impact transmission of an outbreak? Will it take into account misinformation and false narratives around diseases, as well as national and local responses to those diseases?

It was pretty clear with the Ebola outbreak that the solution was not just around vaccines. It was around social anthropology and understanding what was happening on the ground—people and the way they buried people, for example. How does the center work in that respect? This is not like the weather. This is about people and how they are going to change things. I am interested to hear what Caitlin and Dylan think, but also to hear from Jeff and Jim, if they have perspectives on how things have been happening recently or in the past.

RIVERS: It is true that human behavior is one big difference between weather forecasting and disease forecasting. Disease processes are fairly straightforward to study, understand, and codify. But human behavior is a lot more dynamic—it is more difficult to understand and anticipate. Nonetheless, I think incorporating behavior data into our modeling approaches is one of the biggest areas of opportunity, and one of the places where, I think, investment and attention to developing models will be able to yield great results. There is a lot we can do around better understanding: How do people react in the setting of an infectious disease outbreak? What are the ways that we can intervene in those behaviors? How can we translate that into the language of models? I think that this is going to be an important area of growth for the field that we have only just begun to understand and scratch the surface of.

SHAMAN: This is going to be an increasingly important issue. It is certainly possible with wearables and the swell of information that may come from individuals who are volunteering to provide information on some of the micro-scale processes that are really critical for understanding and helping control diseases. The COVID pandemic has made this eminently evident, but we still have not been able to crack it. We know, for instance, that face masks are effective. But we have a very difficult time quantifying the effect of face masks because we cannot observe actual face mask usage—the quality of the mask, whether somebody is wearing them over their nose and mouth or hanging off their chin, or how much of the time an individual effectively wears it out of the house. We have proxies for mobility. We even have Bluetooth engagement, that Apple and Google have come up with, for trying to peg proximity. So there are many interesting devices and there will be means for improving our ability to understand the behaviors that are presenting opportunities for person-to-person transmission, particularly for something as invisible as a respiratory disease transmission event.

As behavior feeds into the transmission process and as we develop better indicators for it—and it does not all have to be technological—this information really does need to be incorporated into these models. It is critical. Getting back to the idea of this difference between projection and forecasting, when you look at what we have been able to do for the COVID-19 pandemic, it really is projection because it is all conditional probabilities. It is "if-then" statements. If society and individuals do this, then this is what we might see in the future.

The problem is that "if" is all wrapped up in behavior. What kind of compliance will there be in social distancing, restrictions on mass gatherings, or usage of face masks? Until we are able to observe that, there is going to be a hole in our ability to project with greater accuracy what will happen in the future with infectious diseases.

FLEMING: I think there is a very close analogy here between pandemic forecasting and weather forecasting. If you look at it historically over a really long period, human expectations have been co-produced with the increasing ability to forecast the weather. I think the examples we are giving are from our current pandemic, but human behavior is malleable, and it does respond to these advances in forecasting. So, you are going to end up changing expectations for the health system as you go.

GEORGE: The only other thing I would add, to the otherwise excellent comments already pointed out, is that the weather analogy is an imperfect analogy. It is not completely apt in all the cases. And we fully appreciate that as well. If there are other analogies that people find more compelling, we would love to have that discussion with folks, too. It is useful for helping think through the components, and for then learning from a different field what major efforts were necessary for moving forward to advanced analytics.

The last thing I would point out though, too, is that there are some encouraging efforts going on right now in the National Science Foundation. They are in the process of scoping out a program to bring in social, behavioral, and economic components, and trying to think through how to augment epidemiological models going forward. I am very encouraged by those efforts. Those are to be determined on how they are rolled out. But there is an explicit agreement among many people in the academic community that we need to understand these things at a much finer granularity going forward, as Jeff was pointing out.

CONNELLY: I know Jim from years when the two of us led a program on the history and future of climate change. My own particular interest is more generally prevision. How do big institutions and organizations try to predict and plan for catastrophic threats? One of the consistent themes in that history is how frequently different forms of prevision are confused and misused. It is sometimes the fault of the experts. But the public, I think in many cases, does not necessarily understand the difference between a forecast versus a projection. We have seen this over and over again in the course of the COVID-19 pandemic.

I would hope that this new center has the public education component. More generally, I hope that more researchers begin to work across different domains, because there are interesting connections and parallels on how to anticipate different kinds of threats. And we can absolutely learn from one another. As historians, we never think that the analogies are going to give us the answers. What they are really for is to raise questions that might not occur to us otherwise. That is really the most useful thing about looking at past efforts and across different domains.

RIVERS: I think the communication challenges that you just raised are one example of why we think it is so important to grow this out into an operational capability. The people building the models might not be the right people to communicate the models. The weather community underwent a very specific and focused effort to understand how best to communicate emergency weather alerts. A lot of work has very specifically gone into understanding how to motivate behavior change. I think we have the opportunity to do that here based on disease forecasts. We are just not in a position to do it right now with mostly researchers at the helm.

HALLIDAY: There is a somewhat related audience question about the politics of data. It is one thing for people to have their own personal views about infections or data, but the politicians and leaders often have reasons to stifle, block, or distort data—starting at sampling and working all the way up the food chain. What does this matter? Of course, it matters hugely because pandemics are global. So, it matters hugely when you are thinking about the international impact of this. Any thoughts about this that go beyond just the personal human aspects to political weaponizing of data?

RIVERS: There are many and various challenges with data, and not all of them are technological. But I would say, one of the best ways to understand the quality of the data you are receiving without any additional context is to have ongoing analysis and to be able to interrogate those data from a lot of different directions—matching it to your expectations, matching it to other jurisdictions. Having in place analytical capabilities to be able to do that regularly I think will drive us to better understand what the problems with our data are.

GEORGE: Having worked in the Obama Administration, where I was able to see the full force of the federal government doing what it could during the Ebola response, and then seeing it essentially on the bench during COVID just broke my heart. That is why, as we were talking about the center, we wanted to figure out how to support the state and local decision-makers. Now, I am not naive enough to think that all mayors and governors are going to follow guidance from the evidence base that the center would generate. But, if we can scale what we have seen in other places to more places that do not have access to it right now, that would be a way of bolstering some of the challenges with politics. It is a thorny issue, and I completely agree with you that we need to think that through more deeply.

SHAMAN: I would also add that it has crept into the weather. We had a hurricane cone of uncertainty that was expanded before the pandemic, for instance, for political reasons as well. We have to work to establish trusted, reliable sources of information. If we can build out a national center and work on some of the communication issues that Caitlin spoke to, then we can build out the trust from the public for the system. It will become harder for it to be weaponized for political purposes. I am also not naive enough to believe that that cannot happen, but these are the efforts that are needed.

I would also add that there are enormous data issues associated with infectious diseases that are perhaps even larger than these political issues. These have to do with the privatization of data that really need to be in the public domain if you are going to make effective, analytic, and informed forecasting decisions. We have enormous amounts of medical records that are privatized and held by medical billing record companies. These records are inaccessible for use by forecasters for doing something for the greater public good, but are bought and sold by companies who gain access to them. A national conversation about this would be welcome to find ways of bringing these data for use in a way that respects people's privacies, but can be leveraged for the public good. This is a very large subject that needs to be addressed as the center is built out.

JAMES: What is also required is to build and find political champions for doing the right thing. That has to be done in a deliberate way. You have to find them, you have to support them, and you have to make sure that they have the ability to speak out. That is really quite important because we must also get the politics right on this. If you get the science and the data right, you also have to get the politics right. I just wanted to add that.

I wanted us to spend a bit of time talking about the national center: What is the plan for building it? What are the steps? What are the stages? There is an audience question about the span of specialties and professions envisioned for this forecasting center. How can current epidemiologists and/or disease ecologists prepare to be part of this? How will the center interface with the CDC and agencies?

I will also add that there is great interest in replicating the center globally. Not built on the same resource base, clearly. If you want to do this in Africa, it is a much more challenged environment. But the idea of a forecasting center and of improving data collection—not only for infectious diseases, but also for weather and other climate-related issues—is of great importance. I think it would be interesting to speak about how the center is going to be built up over the next year, and what the opportunities for application would be globally.

GEORGE: National Security Memorandum 1 calls for the National Security Advisor to plan to establish the center. It resides with the White House to come up with a plan going forward. The dollars are at CDC, so there are resources available. They are going to have to come together with a process to engage external stakeholders. Those discussions on what the appropriate process is to build that out and to move that forward are assuredly happening.

Wilmot made the excellent point that there are other places in the world that have done this fairly well. There are models that we can look to. Hong Kong, the Netherlands, and the United

Kingdom have done this fairly well. We can learn from what has worked well, and from what has not worked well, to try to generate capabilities here in the United States.

Also, there are ongoing efforts. I know that Rick Bright and the Rockefeller Foundation are very interested in trying to develop global early warning capabilities going forward. There is lots of interest in trying to move these efforts in a sustained way. It is somewhat of a non-satisfying answer, because the processes have yet to become apparent, moved forward, and developed. But I am confident that the government is putting those into place so that we can understand how people can actually participate in forming this.

In terms of the specialties, clearly, we are going to need quantitative people, data technologists, and social-behavioral expertise to make the data move, to analyze the data, and then communicate the data. Caitlin also pointed out that describing the nuance and complexity of some of these results into a policy context is not a straightforward thing to do. Just because you can create the model does not mean you can actually explain it to politicians. So, we are going to have to develop that skill set as well. It is going to be a fairly complex endeavor, but I am hopeful given the mandate and the resources that are available right now.

RIVERS: The only thing to add is that the language right now resides with the administration, and the dollars are until expended. We need proper language in the legislative vehicle to institutionalize and permanently keep the center. Annual appropriations would be a boost as well to make sure we do not make this ephemeral—that it is something lasting and enduring.

JAMES: Thank you very much. There is one audience question directed at Jeff: Predicting pandemics—can it really be realized in reality? If so, what are the complications that we face to realize its functionality?

SHAMAN: There are two issues here, and Dylan spoke to this earlier. If we are talking about forecasting diseases that are present in society, that are recurrent, or that have emerged—trying to predict where they will spread, what the burden will be, or how many cases there will be in locations where the pathogen is already present—that kind of forecasting is possible to different degrees, and has been for a generation over the last 10 years operationally. The CDC has been using seasonal flu forecasts. There have been efforts to implement this for dengue and for other diseases as well. It is certainly something that can be done. The accuracy, the reliability, the calibration of those forecasts still need to be improved substantially. And that comes about from all the research that we were talking about in terms of improving the models, improving the observations, and improving understandings of behavior—bringing all that in so that we can provide better information along the way.

On the other side, if you are talking about predicting the emergence of a new pathogen—being able to predict the 2014 Ebola outbreak before it happened or being able to predict the emergence of SARS-CoV-2 before it happened—that issue has really intrigued people for a number of decades, and there has been a considerable investment in it from the federal government. Projects like PREDICT and organizations like DARPA, part of the Department of Defense, have invested considerable amounts of money in such efforts.

The caution there is that it is very difficult to point to an example of predictions made for an event that has never been observed. I do not think you can find it in any field to be perfectly honest. Our ability to predict the weather is because we have observed and understood the weather, the dynamics, the thermodynamics, and the radiative processes of the atmosphere. Even though they are nonlinear, we can build model systems that represent those physics and simulate atmospheric outcomes.

When we do not know an actual process—in this instance because the pathogen is completely novel—it is very difficult to make a prediction. However, there are ways of assessing risk. For instance, there are very interesting ways for understanding reservoirs of disease. What kinds of animals—that we have not seen before, but may be more likely to carry diseases—have the capacity to infect humans with pathogens capable of sustained person-to-person transmission? There are really interesting ways of carrying out such risk assessment. These methods may not be able to tell us that on 3 June 2026, somewhere in Cambodia, there will be a spillover of the next global pandemic, but rather they could provide information such as: Where do we need to do more surveillance? Where do we need to keep an eye on things? Where do we need to improve our response capacity? What are the animal systems and locations that we need to monitor in greater detail?

While I may have presented it in a more skeptical light—in terms of applying the kinds of quantitative methods that we are talking about for forecasting diseases that are out there and among us already—there are other methods that can be brought to bear to help protect us and realize where threats may be coming from.

RIVERS: I absolutely agree with Jeff that the science of predicting disease emergence is probably not to the point where we can rely upon it to give us the earliest of warning systems. But one observation that I have made over the course of years, and that I think Dylan shares, is that early detection is rarely the bottleneck. There is often a long delay between when we realize there is a problem and when we feel inspired to move against it and react. Part of the value, I think, of modeling and analytics is that it can build more confidence in decision makers and the people responsible for pushing go, by helping them to understand: What are the possible future scenarios? How might interventions change the course of those scenarios? I think early action is just as important as early detection.

JAMES: Thank you very much. We are about to close. Before I hand it over back to Matt, I want to ask if Alex had any last comments to make?

HALLIDAY: Thank you. It was brilliant to have such an interesting mixture of people here. Also, it is great to be excited about a future vision—for doing something big in the way that America always has been good at doing. I think it has been great to talk about how to plan for the future in this way.

CONNELLY: Jim, you had your hand up a moment ago?

FLEMING: Since we are in such a long-term effort, as Caitlin mentioned, I think we need multiple seminars and educational opportunities to bring in all kinds of talent. Gro Harlem Brundtland said that we need all of our expertise to address the climate issue. Also, there is the intergenerational aspect—that we will be training and educating new people. So that is a really great opportunity going forward.

CONNELLY: ISERP, as the social science institute at Columbia, is very hopeful, and we have already begun talking with Alex, Jeff, and others across Columbia about how we can support the new Columbia Climate School and how social scientists can join with other scientists in moving that vision forward.

Once again, I want to thank Wilmot James. I want to thank all of our panelists. And I want to thank all of our partners in supporting the series. We have had a lot of help and support from the

Programs in Global Health, the Vagelos College of Physicians and Surgeons, the Earth Institute, the Data Science Institute, and our publishing partner, The Academy of Political Science. Finally, give a special thanks to Harlowe Wang, ISERP's program coordinator, and Loren Morales Kando of the Academy of Political Science. Thank you everyone for joining, and until next time.

WILMOT JAMES: Thank you very much.