United States House of Representatives

Committee on Energy and Commerce

Subcommittee on Oversight and Investigations

Challenges and Opportunities to Investigating the Origins of Pandemics and Other Biological Events

Tom Inglesby, MD

Director, Center for Health Security

Johns Hopkins Bloomberg School of Public Health

February 1, 2023

Chair McMorris Rogers and Subcommittee Chair Griffith, Committee Ranking Member Pallone and Subcommittee Ranking Member Castor, and other members of the Committee, thank you for the chance to speak with you today about the issue of *Challenges and Opportunities to Investigating the Origins of Pandemics and Other Biological Events.* My name is Dr. Tom Inglesby, and I am the Director of the Center for Health Security of the Johns Hopkins Bloomberg School of Public Health and a Professor of Public Health and Medicine at Johns Hopkins University. The opinions expressed herein are my own and do not necessarily reflect the views of The Johns Hopkins University or Health System.

Our Center's mission is to protect people's health from major epidemics and disasters and build resilience. We study the organizations, systems, and tools needed to prepare and respond. Today, I am happy to provide testimony on the challenges and opportunities to investigating the origins of pandemics and other biological events.

High importance of these issues

I commend you for focusing on these issues. Strong capability to determine the source of future pandemics and other biological events may help us diminish the chances of those events. Whether a future pandemic has a natural origin, or is the consequence of an accident, or is the result of a deliberate event, it will be critical to do all that can be done to determine its origin. The term "biological attribution science" has also been used to refer to the body of scientific and investigative work that can be used to determine pandemic (or epidemic) origins, and so I

will also use that term in this testimony. Future pandemic or biological threats could emanate from nature in some form. They could arise from a high consequence laboratory accident. Or they could arise from the deliberate use of a biological weapon from a US adversary. All of these could occur without warning and need to be prepared for broadly. One element of that preparedness is the capacity to attribute the source of future pandemics or other biological events as rapidly as possible.

If after a future biological event, biological attribution identifies, for example, that a certain animal management practice is responsible, that should have major consequences for the future of that practice(s). If it were demonstrated through biological attribution that a pandemic or biological event started in a laboratory, that should have very important ramifications for biosafety – it would be critical to understand how the accident originated and how to reduce the risks of additional events. Biological attribution is also crucial for national security – if a pandemic event or another biological event is deliberately started, then the national security of the country requires us doing all that we can to attribute the source of that. A clearly communicated and strong capacity to attribute the source of a deliberate attack using a biological weapon should also serve as a strong deterrent against that happening in the first place. Not only is it crucial for the US government to understand what initiates a deliberate outbreak, but it is also essential for the US government to avoid being deliberately deceived about the origin. Highly effective bioattribution tools could help not only to identify the accurate source of a biological event but also to debunk a false-flag attempt by adversaries who attempt to falsely assign blame to a country or entity which actually had no responsibility for it.

I also want to call out the importance of attribution science in the effort to grow and power the US bioeconomy. A growing, dynamic sector of the U.S. economy is driven by the products that are made through biotechnology, with every expectation that that portion of the economy here and elsewhere in the world will keep expanding quickly in the time ahead. To keep the US moving fast ahead as a contributor to the domestic and global bioeconomy, we will need to advance attribution science to help us defend the provenance and intellectual property related to U.S. biotechnological products.

The Administration's National Biodefense Strategy -- published in October 2022 - commits to strengthening national attribution capacity (on pages V and X). The strategy says that the US will: "Enhance and sustain U.S. Government characterization capabilities for forensics and attribution, serving the U.S. human, animal, plant, and environmental health and national security communities." It places leadership for this work at the Department of Health and Human Services (HHS), and the Federal Bureau of Investigation (FBI), with support from the Department of the Interior (DOI), the U.S. Department of Agriculture (USDA), the Department of Energy (DOE), the Department of Homeland Security (DHS), the Environmental Protection Agency (EPA), the Centers for Disease Control (CDC), and the Intelligence Community (IC). The National Biodefense Strategy also says that the U.S. government will: "Strengthen the capability of the UN Secretary-General's Mechanism for Investigation of Alleged Use of Chemical and

Biological Weapons (UNSGM) to determine the facts, including attribution, regarding the alleged use of biological or toxin weapons." The strategy says that the lead for this work will be the Department of State (DoS), the Department of Defense (DoD), and the FBI.

The range of agencies listed in support of these mission show how multidisciplinary this work must be to succeed, and how distributed the necessary talents and resources are in the U.S. government. Biological attribution requires the capacity for strong, rapid domestic response, as well as the ability to work internationally with other partners with the goal of determining the origin of major biological events.

Important GAO Report analysis and recommendations

The GAO report **Pandemic Origins: Technologies and Challenges for Biological Investigations** provides a highly valuable assessment of the approaches used for biological attribution as well as policy recommendations aimed at advancing national capability in the field.

Biological attribution science relies on a combination of sources and approaches. These include genetic sequencing and bioinformatics and access to genetic databases. As the GAO report describes, the purpose of this kind of sequence analysis is to compare one pathogen's sequence to the sequence of other pathogens in an effort to find matches or pathogens that are more closely related than others. This work, which includes the study of the evolutionary history and relationships among or within groups of organisms, is called phylogenetics and is a critical field of science in biological attribution. The tools that are now available for genetic sequencing are extraordinary and sophisticated. But for sequencing to be of greatest value, access to samples, genetic sequences, and genetic databases will be critical. Data from some areas of the world (e.g., the U.S. and U.K.) are overrepresented in some large genetic databases, with no data included from other parts.

Another major weakness in approaches to biological attribution is that while the tools of genetic sequencing and bioinformatics are quite powerful and accurate, they cannot always reliably distinguish whether a pathogen came from a natural source vs whether it was passaged, edited or engineered in a laboratory. More work remains to develop bioattribution tools that can address critical questions like these.

For instance, other areas of science, including the study of proteins (proteomics), the study of sugars on proteins (glycomics), and the study of environmental factors resulting in modifications of genetic materials (epigenetics) have great potential to provide value to an origins investigation, but they are not yet fully developed in ways that make them central approaches to this work.

The success of bioattribution scientific work may also depend on access to public health surveillance data, animal surveillance data, environmental data, and human clinical data. Epidemiological data about patterns of spread and risk factors for those infected may be pivotal. And all of this data would be critical as early as possible in a novel pandemic or after a

biological event, as close to the initial case or cases as possible, before the data is lost to time, degraded or just no longer possible to obtain. Some human or animal data, like serology data, may not be possible to collect for a substantial period of time because of the time it will take for researchers to identify rigorous serological markers. Field work in collecting the right samples requires people with the necessary skills, materials, technical approaches, and again requires access to the places where samples need to be collected. Biological attribution work will also depend on an investigative process with the highest possible standards, great integrity and trust.

Specifically on the issues of data sharing domestically, it will be important in future events for states to share their clinical and public health data with HHS and CDC so that the federal government can have an integrated picture of what is happening, trends, patterns that could help identify sources. That authority to collect data from around the country and integrate it to develop a full national picture is something that HHS/CDC are seeking.

The GAO report provides concrete findings and recommendations to improve U.S. capabilities, all of which I support. A few brief comments on the report's recommendations:

- **Establish multilateral agreements for accessing and sharing samples and genetic sequence data**

Other than for influenza, we don't yet have international agreements in place to share genetic sequence information at the start of a new epidemic or pandemic. Countries may or may not share that information in the course of a new pandemic.

Multilateral or global agreements for sharing samples, sequence information, bioinformatic information and the other sources of critical data above would be a major step forward. International organizations like the World Health Organization could help to develop those agreements, but it will not be easy work. Some countries have expressed concern about inequitable access to medicines and vaccines that are derived from biological samples taken from their territories. Data sharing agreements would likely need to address those issues and provide incentives for rapid sharing of data and information. HHS Office of Global Affairs and the DoS should make this a priority in their international engagement work, and the U.S. should be a leader in encouraging the implementation of these kind of data sharing agreements.

- **Develop standard processes for genetic sequence database use**

Different genetic sequence databases use different submission tools for sequence information, different processes, different user interfaces, different standards. There are errors in some databases that are difficult to correct and may become permanent. Metadata (time, place of collection etc.) that is provided with sequence information is too vague, variable or missing in places.

To deal with this, federal policymakers, in partnerships with scientists, database administrators and other key stakeholders, should try to drive toward more efficient, standard approaches to database submission, storage, management, metadata requirements, sharing of information, etc. Ultimately since these databases are located and used internationally, this work will also be inherently international and will require close collaboration with the developers and owners or other critical databases around the world.

This work, too, will not be easy or rapid, given the broad distribution of databases, and the widely varying approaches, different institutional approaches and controls, costs etc.

- **Improve current or develop new, genetic sequence database tools**

The GAO report rightly calls out for investment in and advancement of new database tools and user interfaces. This is particularly important because of the anticipated rapid expansion of genetic sequence data in the time ahead.

- **Encourage the development, retention and growth of a workforce with the critical skills needed for pandemic origin investigations**

As the GAO concludes, many skillsets are critical for this work – biology, virology, microbiology, immunology, epidemiology, ecology, genomics, bioinformatics and computer science to name just some. This is inherently multidisciplinary work.

Providing incentives, fellowships, training opportunities to support these fields generally, and to have at least some specialized attention to the study of pandemic origins and biological attribution science will be important. Clearly, this is a long process, but having a strong workforce in these fields will be important in developing and maintaining the right kind of expertise in the country for this mission. Investing in the development of the workforce for this mission would not only be important for the success of this bioattribution work, but would also be valuable in strengthening STEM and biotech capacity of the U.S. and its bioeconomy.

- **Augment or develop a national strategy to better coordinate and collaborate domestically and internationally on pandemic origin investigations**

A commitment to improving biological attribution is in the recently released National Biodefense Strategy. But the GAO rightly calls out that the US does not yet have an articulated strategy for developing the strongest possible national capacity for biological attribution or investigating pandemic origins and other biological events. Having such a strategy, with assignments, milestones, timelines and budgets attached to them would drive national capability forward.

Recent White House Office of Science and Technology Policy bioattribution meeting

In December 2022, the White House Office of Science and Technology Policy convened a group of scientists, policy focused researchers and government officials to assess current national efforts around biological attribution science. A summary of that meeting was prepared by my colleagues and can be found on line at: https://www.centerforhealthsecurity.org/our-work/publications/discussion-on-the-future-science-and-technology-of-biological-attribution

Meeting discussion and take-aways were generally in line with the major points made in the GAO report. However, there were some additional insights provided and recommendations made over the course of the discussion.

- **Genetic Engineering Challenge:** A recently completed and published Genetic Engineering Attribution Challenge (https://pubmed.ncbi.nlm.nih.gov/36450726/) was proposed in the OSTP meeting as one model for driving one component of biological attribution science forward. Colleagues at my Center were part of the team that ran that competition. Teams used different machine learning (ML) approaches to formulate predictions regarding the lab that specific genetically engineered sequences came from, based on sequences that had been previously deposited by those labs in a publicly available database. The winning teams were able to substantially outperform past machine learning approaches that took on this challenge. Not only were these models better at identifying where sequences came from, but they were better at excluding labs where such sequences might have been produced (called negative attribution). This was a relatively small, foundation funded study – larger competitions might drive faster approaches. These kinds of ML approaches would be useful under specific conditions: if a lab is the source of a future pandemic or biological event, this kind of approach would work in identifying a lab source if that lab had previously published its sequences or deposited them in a public database.

- **Gaps in data:** There was substantial emphasis in the OSTP meeting on the paucity of major databases, and on the recognition that, even if more sample collection is funded and supported there will be major gaps in data. Bioattribution science will need to find ways to move ahead despite that challenge. Given their important to future database management and construction, it was advised that responsible officials from NIST and NCBI should be central to future biological attribution discussion and planning. There was also a strong focus on the importance of standardized approaches to data collection and metadata practices. It was emphasized in the meeting that attributing the source of future biological events may require incorporating many different kinds of data – e.g. pathogen characteristics, genetic sequence, evolutionary data, epidemiologic data and others – so anticipating the need for all of this and the need to integrate it in an investigation should be an operating assumption in building this capacity.

- **Federal investment:** The meeting also identified areas in need of additional national investment and attention that could strengthen the U.S. biological attribution capacity, including proteomics, database development, machine learning tools, and new approaches to sharing data from different sources that do not require disclosing proprietary information.

Forward progress and recommendations

Strengthening biological attribution science will strengthen our national capacity to identify the origins of future pandemics, biological attacks, high-consequence laboratory accidents, or other significant biological events. It is important as a tool for understanding whether a threat emanates from nature, or through an accident, or through deliberate misuse. This is critical for public health and national security. Being able to determine who carried out the deliberate use of a biological weapons could serve as a deterrent to those who might consider using biological weapons against the U.S. It could also help to identify efforts of those who might attempt to hide attribution of an event and blame it on another entity (false flag events).

Biological attribution capacity is likely to also be very important, and will become increasingly so, for securing the many elements that are produced by the bioeconomy, whether they are crops, foods, energy, medicines, or other materials or products that biology is helping to create.

Making forward progress on biological attribution will not only be very important for the U.S., it will also be good for all governments. It is in the interest of the international community to understand the source of future biological events, whether natural, accidental or deliberate. The sooner we can understand what caused an event, the faster the international community can work together to diminish the chances it will happen again. And if countries have a strong sense that future manmade events will be attributed to specific countries, or even to laboratories within them, they will be more likely to ensure strong biosafety, biosecurity, and governance systems, and far less likely to consider the use of biological weapons.

Fortunately, strengthening biological attribution science is called out as a priority in the National Biodefense Strategy. The GAO has written a valuable report and provided a series of good recommendations. These developments combined with the interest of this committee in strengthening biological attribution as evidenced by this Hearing, I think suggest there are opportunities for important progress. Substantive progress will require priority setting, financial investment, and sustained collaboration—both domestically and internationally.

In terms of recommendations to this committee, I will start by supporting moving forward on the options that GAO has set out in its report, for reasons noted above. I also have these additional recommendations:

- **Identify lead agencies and where major responsibilities for this work reside:** A key component in developing a national strategy for investigating the origins of pandemics and other biological events should include the identification of the lead agency/program

office(s) in the USG that have responsibility for biological attribution. The National Biodefense Strategy has identified a number of lead offices, and it may be that different offices must lead different elements of building this capability given that there are different critical components of this work that require different skillsets and operations (e.g. research into new tools and scientific approaches; potential new prize grants to drive the field work; machine learning research relevant to this field; operational preparedness to do rapid sequencing, other scientific studies that would contribute, database investigation at the very start of a new biological event; preparedness for investigative work that would stand up in domestic or international settings; preparedness for international collaboration during investigations, etc.).

If it is the case that leadership for biological attribution is distributed to different agencies, a valuable step now is identifying which specific biological attribution responsibilities reside in each. For example, what is the responsibility of HHS vs FBI? Within HHS and FBI, which agency/program is responsible for the work? Similar questions of responsibility apply to the DOD and DOS on international components of this work. And for the other agencies that are listed as in support of this work in the National Biodefense Strategy. As with other national responsibilities that are distributed in the U.S. government, or "whole of government", strong White House interagency management and oversight will remain important.

- **Support scientific research and tool development:** As with most scientific capabilities that underpin a national priority, much of the new science that will occur to move the field forward will reside outside the government, either in research universities or the private sector. Funding, convening and building those external communities of practice that are working toward a better national capacity will be important.

- **Budget Planning**: Another key component for setting the right strategy is to define resource needs for this work. For those U.S. government agencies/offices with responsibility for leading and implementing this work, what are their budgets related to this work now? While the USG does provide some internal and external funding for this work, it's not clear what the budget for this work is currently, or what the budget needs are overall. Should additional funds for critical areas of research, preparedness and operations related work be authorized and appropriated? My intuition is that the answer is likely yes, but it is good that Congress is now seeking more information on this so it will be able to determine what more is needed.

  Strong funding for this work will be important to strengthen the science, the international partnerships, the workforce and the government organization of this work. But in addition, funding is also crucial in allowing the government to do exercises to test its capacities and operations around this work. If the government does not regularly exercise how it will

pursue pandemic origin investigations and attribute the source of future biological events, then it will not have a functioning process in place for future events.

- **Sharing Data:** Congress should also ensure that federal agencies involved in biological attribution science are sharing and combining datasets that are relevant. This is a strategic capability that requires collaboration between different parts of government, and may require a congressional mandate to make such sharing a reality rather than an ideal.

In conclusion, the country does not yet have in place the scientific knowledge, operational plans, investigative process or international partnership that can reliably and assuredly identify the origins of future pandemic and biological events. Building stronger national capacity around the power to identify the origins of future pandemic and biological events, also called attribution science, should be a high priority for the Administration and Congress. I am very pleased to see this committee taking on this issue. As we make forward progress in building this capability, we will be better prepared to investigate future biological events, better able to understand whether events emanated from nature or from laboratory settings, more likely to find where we might need changes in biosafety, better able to deter deliberate biological threats, and better able to protect our U.S. bioeconomy. Thank you for the opportunity to speak to you today and I would be pleased to answer any questions you may have.