Testimony of David A. Relman, MD

Thomas C. and Joan M. Merigan Professor, Departments of Medicine, and of Microbiology & Immunology, Stanford University School of Medicine; Senior Fellow, Center for International Security and Cooperation, Stanford University, Stanford, CA

Before the U.S. House of Representatives
Committee on Science, Space, and Technology
Subcommittee on Investigations & Oversight
Hearing on “Principles for Outbreak Investigation: COVID-19 and Future Infectious Diseases”
Wednesday, July 14, 2021

Chairman Foster, Ranking Member Obernolte, and members of the Subcommittee, thank you for the opportunity to join you today for this important hearing and discussion about principles for outbreak investigation. My name is Dr. David Relman. I am a professor of Medicine, and of Microbiology & Immunology at Stanford University School of Medicine, and a senior fellow at the Center for International Security and Cooperation, also at Stanford University. I am also Chief of Infectious Diseases at the Veterans Affairs Palo Alto Health Care System, but am appearing here today on my own time. Today’s hearing takes place at a critical and consequential moment for this nation and the world, as we struggle to address one of the most calamitous public events of the past century, COVID-19.

I will begin with some brief remarks about the emergence of infectious diseases and the changing factors that underlie recent outbreaks, as well as create risks for future ones. I will present some of the hypotheses that could explain the origins of the current pandemic, including the evidence that supports these hypotheses, and summarize where we stand with respect to an understanding of pandemic and virus origins. I will then discuss the key elements of pandemic investigations, present and future, offer thoughts about why it is so important that we seek greater clarity about the origins of outbreaks and pandemics, lessons learned so far in assessing COVID-19 origins, and suggest some next steps. Some of my thoughts on this topic have been published elsewhere¹.

The opinions expressed in my testimony are my own and do not represent the views of my employers, Stanford University and the U.S. Department of Veterans Affairs.

Emergence

Ever since the time that humans first congregated in close proximity, domesticated animals, and cultivated plants as food, microbes have had increasingly important impacts on

morbidity and mortality. Human history was regularly punctuated by plagues and large-scale die-offs until the advent of public health measures such as isolation and quarantine in the 17th century, sanitation in the 18th century, and vaccines in the 19th century. In the second half of the 20th century, methods for microbial discovery and diagnosis became more powerful, and revealed microbes that were new causes of human, animal, and plant disease, along with others that had caused disease for some time but were unrecognized. By the early part of this century, a rapidly growing number of infectious disease outbreaks suggested that microbial disease “emergence” deserved special attention. In 1992, a study at the Institute of Medicine (now, National Academy of Medicine) organized the factors responsible for emerging infections according to three categories: those related to microbe, host, and environment. Two environmental factors appeared to be unusually important, both of them anthropogenic: climate change and changing land use leading to more frequent human encounters with animals harboring zoonotic pathogens. In 2003, a re-assessment of emergence by the same group at the Institute of Medicine (on which I later served for a dozen years) identified the previously described factors, as well as social, political and economic factors, and the growing human capabilities for genetically engineering microbes and hosts. This work serves as an important backdrop to current thinking about the origins of the ongoing pandemic.

Hypotheses about the origins of COVID-19

The initial characterization of the causative agent of COVID-19, SARS-CoV-2 revealed it to be a beta-coronavirus whose closest known virus relatives had been found in bats in China. Based on this finding and the known capacity of bats to serve as reservoirs for coronaviruses, it is assumed that at some point in the past, SARS-CoV-2 colonized and evolved in bats. How, when, and where SARS-CoV-2 first encountered and infected a human, and then started a chain of uncontrolled human-to-human transmission, is the key question.

Hypotheses about the origins of COVID-19 and the circumstances of that first encounter can be grouped into two major categories: 1) those involving a natural “spill-over” (from natural host to accidental host) either directly from a bat to a human, or indirectly through an intermediate host animal; and 2) those involving a spill-over facilitated by research activities at a laboratory. This second category includes collection of animal samples by researchers and the handling and study of these samples and their viruses in a laboratory. In both cases, the initial fateful encounter may have taken place in the natural habitat of the animal host when a human intruded into this environment, e.g., a cave, or it may have taken place closer to the places where humans live and work, when humans deliberately relocated the animal hosts to these sites, e.g., a city market or a laboratory. In both of these hypothetical categories, the actions of humans may be responsible to varying degrees for precipitating the pandemic. The difference between the two

---

sets of hypotheses concerns the kinds of human activities implicated in the origin of the pandemic: in the first case, violation of natural habitat boundaries and the promulgation of wildlife trade; in the second case, the manipulation of potential pathogens and the strategies used to understand microbial threats.

**Evidence in support of COVID-19 origin hypotheses**

Many have gazed at the genome sequence of SARS-CoV-2 and from this sequence, tried to infer the origin of the virus. Despite much analysis and discussion, the sequence yields only limited insight. The basic problem is that the immediate, most recent viral ancestors are missing (are not available in the public domain). Therefore, the most recent evolutionary history of SARS-CoV-2, and the circumstances under which this most recent evolution took place, is unknown.

It is reasonably clear from available coronavirus sequences that ancestors of SARS-CoV-2 resided in bats. The most closely related known virus, RaTG13 was apparently detected and recovered from a sample collected from a bat in a mine in Yunnan province in southwestern China. Yet, RaTG13 is not that closely related; it is estimated to have shared a most recent common ancestor with SARS-CoV-2 several decades ago. RaTG13 is an unlikely progenitor of a hypothetical laboratory-based genome engineering effort that would have produced SARS-CoV-2. Analyses based on all available genome sequences of clinical SARS-CoV-2 isolates have produced estimates of the timing of the first human cases of COVID-19 within the period beginning mid-October and ending late November 2019.

Controversy has arisen over the seemingly unusual sequence-based and functional features of SARS-CoV-2, such as its furin cleavage site or its apparent high degree of adaptation to human cells and receptors, which, to some observers, indicate a laboratory origin of the virus. To others, these features are best explained by natural origins and suggest to them that the virus could not possibly have originated in a laboratory. However, both conclusions rely on unfounded assumptions, and lack direct evidentiary support. The missing, most recent viral ancestors of SARS-CoV-2 may well turn out to be found somewhere in the natural world, especially because efforts to sample the immense diversity of coronaviruses, e.g., in bats, so far are woefully incomplete. On the other hand, these ancestors may turn out to be found in the large collections of uncharacterized samples located in laboratories that have dedicated themselves to the study of these viruses. They may have accidentally infected unknowing laboratory personnel either at the time of sample collection or back in the laboratory during subsequent efforts to cultivate or manipulate viruses, or during disposal of samples afterwards. Many of the steps undertaken in laboratories to study animal samples and associated viruses are not closely monitored using deep sequencing methods. Importantly, without direct detection of SARS-CoV-2, either in a natural environment or in a laboratory, all we have is speculation and hunches.

---


So, where does this leave us?

There are four kinds of circumstantial evidence that favor the first set of hypotheses (natural origin):
1. **History.** Because nearly all previous infectious diseases outbreaks involving a pathogen associated with animals are presumed to have begun with a natural spillover event\(^9\),\(^10\),\(^11\), one might assume that the same kind of event caused this pandemic.
2. **Immense viral diversity in nature.** Recent efforts to catalog viral diversity in bats in China indicates that much of this diversity has yet to be discovered\(^12\). One might then logically assume that SARS-CoV-2 resides in nature but that surveillance studies have been so far insufficient to detect it.
3. **Extensive wildlife trade.** Despite the known risks, humans continue to collect, raise, and distribute live animals that are potential sources of viral pathogens, or their derivative parts, to markets where encounters with humans are likely\(^13\).
4. **Natural spillovers happen more often than we had thought.** Surveys of humans living near potential animal hosts of zoonotic pathogens have revealed evidence of past asymptomatic and therefore unrecognized infection\(^14\). Such evidence was published by investigators at the Wuhan Institute of Virology (WIV) after a blood-based survey for antibodies indicating past exposure to bat-associated coronaviruses in four villages near the mine where RaTG13 was found\(^15\).

There are also four kinds of circumstantial evidence that favor the second set of hypotheses (laboratory-associated origin):
1. **Geography.** The closest known virus relatives of SARS-CoV-2 were found in bats in Yunnan province more than 1000 miles from the city where COVID-19 appears to have first emerged, Wuhan. Surveys of bats across China suggest that beta-coronaviruses are found in greatest number and diversity in Yunnan and other locations in southern China, rather than further north\(^16\). The survey for exposure to bat-associated coronaviruses mentioned above included 1000

---


blood samples from Wuhan; all were negative\textsuperscript{17}. At the same time, the laboratory with one of the largest known collections of bat samples and bat-associated coronaviruses (WIV) is located within the city of Wuhan.

2. \textbf{Known work at laboratories in Wuhan.} WIV not only contains one of the largest collections of bat samples and bat-associated coronaviruses (the WHO Report from March 2021\textsuperscript{18} mentions more than 19,000 bat samples, 13\% of which were said to be positive for coronaviruses), but their published work indicates efforts to cultivate viruses from bat samples, as well as recombinant engineering work that created hybrid coronaviruses from a variety of poorly characterized bat-associated viruses\textsuperscript{19}. It is not necessarily always clear to researchers when experiments produce limited growth of a virus or limited growth of a recombinant that arises from multiple co-existing viruses in an animal sample. Risk of inadvertent infection of a researcher is especially high when growth is not recognized. With respect to the virus engineering work, none of the published hybrid viruses nor their source genomes are plausible recent ancestors of SARS-CoV-2; however, this kind of work does pose risks. WIV and other laboratories around the world working with coronaviruses have not provided a comprehensive description of recombinant engineering experiments that have been undertaken in recent years, nor a comprehensive list of virus sequences generated at their facilities.

3. \textbf{Laboratory accidents happen more often than we had thought.} A surprisingly large number of accidents involving dangerous pathogens (so called, Select Agents) has occurred in the United States during the past decade, even at some laboratories with the most stringent safety conditions\textsuperscript{20}. In 2003-2004, three separate laboratory accidents in three countries resulted in infection of personnel with SARS, including the death of a family member\textsuperscript{21}.

4. \textbf{Lack of transparency and incomplete data.} A failure to alert the world about six cases of SARS-like illness in 2012 and to provide a full report of the laboratory investigation of them\textsuperscript{22}, as well as the incomplete or delayed information about work that has been undertaken on bat-associated coronavirus diversity\textsuperscript{23}, suggest that the research community has not yet provided all of the information available that might contribute to a more complete understanding of the origins of this virus and pandemic.

Again, all four kinds of evidence for each of these two sets of hypotheses are just circumstantial. SARS-CoV-2 has not been found anywhere other than in infected humans, and in


\textsuperscript{23} Bloom JD. Recovery of deleted deep sequencing data sheds more light on the early Wuhan SARS-CoV-2 epidemic. bioRxiv 2021.06.18.449051; doi: https://doi.org/10.1101/2021.06.18.449051
animals that have been infected by humans or on fomites contaminated by infected humans. Neither set of hypotheses can be ruled in, nor ruled out.

I have not and will not offer a personal opinion about which hypothesis I believe is more likely, because I do not think that opinions at this point, in the absence of additional primary, verifiable data, are useful, especially in the present political climate.

**Goals and key elements of outbreak investigations**

The goals of an outbreak investigation are to provide greater clarity about the circumstances and factors that contributed to the event. A definitive answer is often not achieved, nor is it necessary for an investigation to yield important benefits. First and foremost, greater clarity about the circumstances and contributory factors that underlie the origin of an outbreak will enable us to anticipate and prevent future outbreaks; for example, targeted interventions and early warning strategies can be designed and tested based on an understanding of mechanisms and risks. Second, every investigation, if undertaken thoughtfully, provides an opportunity to learn how to undertake the next investigation in a more effective manner. Third, any reduction in uncertainty about the origins of a pandemic can help minimize further recriminations, speculation, and conflict.

Some have argued that further clarity about the origins of the current pandemic is not needed, because we already know enough to motivate efforts to address both sets of hypotheses. I disagree. As I have written elsewhere [24], if scientists are not forced to confront the issues of laboratory safety and risky research in a serious manner, history suggests that we will not do so. I will set aside the semantic issue of what defines “gain-of-function” research. More importantly, we need to consider whether there is life sciences research that is so risky that it should not be undertaken. I view the work performed at WIV and described in Hu et al. [25] to be unusually risky. The strategy they chose to characterize novel coronaviruses detected in bat samples relied on creating hybrid viruses in the laboratory starting with a previously characterized bat-associated coronavirus, called WIV1, that had been viewed by experts as already “poised for human emergence” [26] and then inserting the all-important spike protein from the novel viruses. Although the purpose was not to make a virus with greater capabilities for human infection, this could easily have been the outcome. If we learn nothing more about the work undertaken around the world to explore the properties of novel coronaviruses, life scientists and funding agencies are likely to forego the difficult discussions that are needed about risk assessment, risk mitigation, and ‘red lines’.

Some of the key elements of a credible investigation concern process, others concern data. On the issue of process, there are at least six elements worthy of consideration. 1) Objectivity--A priori assumptions should be specified and challenged. Findings should be based on verifiable data, are useful, especially in the present political climate.

---

24 Relman DA. “Lab leaks happen, and not just in China. We need to take them seriously”. Washington Post, June 2, 2021. [https://www.washingtonpost.com/outlook/2021/06/02/lab-leak-investigate-risky-research/](https://www.washingtonpost.com/outlook/2021/06/02/lab-leak-investigate-risky-research/)


on verifiable data and should be reproducible. The integrity of any scientific process must be assured. 2) Transparency—The process of creating the terms of the investigation should be shared with the public in an ongoing manner. Primary data, methods, and the basis for any conclusions should all be placed in the public domain. 3) Representativeness—The interests of all relevant key international constituencies should be represented in the planning and execution of the investigation. The latter should be international, and include nations with diverse levels of resource, and differing kinds of impact by the pandemic. 4) Expertise—Relevant scientific expertise is crucial but should not be defined too narrowly. In the case of COVID-19, expertise must extend beyond virology, infectious diseases, and public health, to include ecology, evolutionary biology, molecular biology, comparative medicine, sociology, political science and security, forensic science, and international law. 5) Independence—The conduct of the investigation must be independent of political interference. In addition, the investigation must be subject to independent oversight and review. 6) Conflicts of interest—Any nation suspected of serving as the place of origin of an international outbreak, any organization with responsibilities for preventing or managing an outbreak, and any scientist whose work is suspected of playing a role in the origins of an outbreak, will all have self-interests that may impede an effective and objective investigation. These conflicts must be articulated and responsibly managed.

With respect to data and materials needed for an effective investigation of a zoonotic outbreak, there are several categories to be considered. 1) Archived animal samples and derivative data—Such samples will be derived from animals in their natural habitats, from various points along the path of wildlife trafficking, and at the places of suspected human encounter where infection was first detected. 2) Archived clinical samples and derivative data—These samples and data will include those from suspected early cases, public health surveillance, wastewater, and social media. 3) Laboratory data—From any laboratory that may have played a role in the origins of an outbreak, a variety of records, data, and samples will be important. These should include records of field activities, animal sample inventories, virus isolate collections, sequence databases, experiment logs and methods, records of all genetic constructions, safety records, and personnel health records and clinical specimens.

Controls are critical for any investigation and should include populations and hosts clinically unaffected early in the outbreak that share features with those affected. The veracity and provenance of samples and data must be addressed. For example, widespread dissemination of a pandemic virus will lead to contamination of a wide variety of materials and spill-back into previously uninfected animals. Thus, efforts to assure the integrity and provenance of any archived sample are essential.

**Observations about the current state of affairs regarding the study of COVID-19 origins, and next steps**

While it is unrealistic to believe that the origins of a pandemic should be entirely understood one and a half years later, we should expect by this point a more coherent, balanced, coordinated, and international effort to investigate COVID-19, and a more substantial body of verified data in the public domain. Instead, the origin issue has become grossly politicized, opinion valued more than rigorous analysis, and a balanced scientific discourse hard to find. Too many good scientists have failed to challenge their assumptions or provide proper scientific
support for their conclusions. Some have formulated their opinions based on subjective assessments of likelihoods from prior outbreaks. There has been ‘group-think’ on all sides of this issue. Conflicts of interest have not been properly addressed. Many scientists are hesitant about addressing this issue because of fear that their comments will be misinterpreted. All of this threatens to harm the credibility of scientists and of the scientific enterprise.

In the midst of this situation, there are positive signs of useful engagement and discussion, a willingness by some scientists and policy-makers to deliberate on paths forward, and some first steps. I helped to organize a letter by a group of 17 other scientists directed at our colleagues and published in a scientific journal that called for a sober and dispassionate assessment of what we know and don’t know about COVID-19 origins, and for a renewed process that relies on sound principles of scientific investigation.27 We did not express an opinion about the source of the pandemic, and suggested that scientists help lead in this effort relying on principles that we know well. The presidents of the U.S. National Academies of Science, Engineering, and Medicine have also issued a similar call28. I believe that national academies of science and other credible organizations that represent the global science community have an opportunity and some responsibility to participate in investigations of pandemic origins. The US NASEM has a productive history of bilateral ‘track 2’ engagements with their counterparts in China, Russia, and many other countries; it is this kind of partnership that will be most effective now. However, groups of scientists will need to proceed under the auspices and authorities of those that represent other global communities, including national governments.

Discussion is underway in many quarters to examine whether there is an effective international governance mechanism for leading an investigation of COVID-19 origins. The World Health Organization would have been an obvious candidate, but during the past year, it has created some doubt as to its ability to operate without undue degrees of external influence and with sufficient degrees of impartiality. Alternatives are well-worth exploring. And although I have some bias and conflict, given my participation in these efforts, I find useful proposals that have arisen from the ongoing work of the COVID Commission Planning Group (Miller Center Foundation, University of Virginia)29 and work by the Nuclear Threat Initiative and World Economic Forum who have proposed a new global entity30 to fill gaps in the current global biosecurity governance scheme and to provide a mechanism for international investigation of outbreaks.

---

In closing, as I have written previously\textsuperscript{31}, a more complete understanding of the origins of COVID-19 and any other international health emergency clearly serves the interests of every person in every country on this planet. Despite the challenges, we must remember that delineating the pandemic origin story will help elucidate the very nature of our precarious coexistence within the biosphere.

I will be delighted to entertain any questions.

\textsuperscript{31} Relman DA. Opinion: To stop the next pandemic we need to unravel the origins of COVID-19. Proc Natl Acad Sci USA 117:29246-48, 2020. \url{https://www.pnas.org/content/117/47/29246}