

COMMENT

Why the world needs more transparency on the origins of novel pathogens

Marietjie Venter on behalf of the WHO Scientific Advisory Group for the Origins of Novel Pathogens

Collaboration and openness are essential to minimize the risks of future pandemics, says the World Health Organization's scientific advisory group, SAGO.



Researchers catch bats emerging from a cave in Thailand to test them for coronaviruses. Credit: Adam Dean/New York Times/Redux/eyevine

On 5 May, the director-general of the World Health Organization (WHO), Tedros Adhanom Ghebreyesus, declared that COVID-19 is no longer a public-health emergency of international concern. This signalled the end of the acute phase of a pandemic that has caused almost 7 million confirmed deaths worldwide. But more than three years after COVID-19 was declared a pandemic, where the coronavirus SARS-CoV-2 came from, and how it first came to infect humans, remains unclear.

Understanding the origins of novel pathogens that could result in major outbreaks, epidemics or pandemics is essential to prepare for the next emerging disease. In recognition of this, the WHO established the Scientific Advisory Group for the Origins of Novel Pathogens (SAGO) in 2021. The 27 members of our group, from 27 countries, include specialists in epidemiology, virology, human and animal infectious diseases, ecology, genomics, and biosafety and biosecurity.

Since SAGO was established, we have evaluated what new data are available to aid investigation of the origin of SARS-CoV-2, met with scientists from around the world and made recommendations about the kinds of data and studies that still need to be collected or pursued. Another key task has been to develop the WHO Global Framework to define and guide studies into the origins of any emerging or re-emerging pathogens of epidemic and pandemic potential. This document is currently under review and due to be published later this year. It is the first attempt to bring together a broad suite of experiences from investigations into the origins of multiple high-risk pathogens, including SARS-CoV-2, mpox (an infectious disease caused by the monkeypox virus), novel influenza viruses, Marburg and Ebola.

The effectiveness of our ongoing efforts to guide investigations into the origins of high-risk pathogens, however, depends on robust science following key principles of outbreak investigation, as well as collaboration, transparency and trust among all stakeholders. To minimize the risk of further disruptive outbreaks, and their associated mortality and burdens on health and economics, various groups must come together — including scientists, public-health officials, governments and global health agencies.

Data drive

In March 2021, together with Chinese researchers, an international team of experts convened by the WHO produced an initial joint report evaluating four possible ways in which SARS-CoV-2 might have emerged¹. Since November 2021, our advisory group has met numerous times and assessed what studies and data relevant to the origin of SARS-CoV-2 have become available since that 2021 report, and to determine what other studies are needed.

We have assessed investigations of human cases of respiratory illness in China in the months leading up to the first recognized transmission of SARS-CoV-2 in December 2019. We have explored studies of animals that are susceptible to SARS-CoV-2 or closely related viruses, and cases of reverse zoonosis, in which the virus has jumped from humans to animals. We have examined environmental studies — involving samples and swabs taken from the Huanan Seafood Market in Wuhan, China, which was the early epicentre of the COVID-19 pandemic, and wastewater samples collected later from all over the world. Lastly, we have reviewed investigations of the genomics and evolution of the virus, as well as biosafety and biosecurity considerations.



A stall selling food products at a market in Wuhan, China, in January 2021. Credit: Hector Retamal/AFP/Getty

Our ongoing review of the available evidence has made it clear that more work is needed. In June 2022, we published our preliminary report², which set out what studies still need to be conducted to understand how the COVID-19 pandemic began. In it, we urged China and other nations to investigate a range of samples that had been collected in the months leading up to the first known cases in December 2019. These include samples taken in research or clinical settings, and as part of environmental surveillance efforts, such as wastewater samples.

We encouraged governments to analyse data from surveillance samples for influenza and severe acute respiratory infections, as well as serology data (from blood banks, for instance) collected during the months before December 2019. The presence of specific antibodies in blood at this time could indicate a previous infection with SARS-CoV-2, which would, in turn, signal that there had been unnoticed circulation of this new pathogen. We also urged China to investigate animal samples taken as far back as 2003 (when SARS-CoV, the coronavirus that causes severe acute respiratory syndrome, was first detected) from species that can be infected with SARS-like coronaviruses, including bats from across southeast Asia.

We requested that researchers around the world make any sequencing data from SARS-CoV-2-like viruses available on global genomic databases. We strongly recommended that researchers in China investigate the upstream sources of the animals and animal products that were present at the Huanan Seafood Market before it was closed on 1 January 2020. We also

requested access to data or documentation on any possible breaches in biocontainment, or on occupational hazards among lab workers or field investigators from China and other countries working on SARS-like viruses. We have not yet received any new data.

Besides the work on SARS-CoV-2 origins, SAGO has been working with the WHO to develop a global framework to define and guide studies into the origins of emerging and re-emerging pathogens of epidemic and pandemic potential. This framework lays out six areas of work that should be pursued were any pathogen with epidemic or pandemic potential to emerge or re-emerge in a region where it has not been found before.

These work areas are:

1. Early investigations to establish what activities and factors could have resulted in the initial infections, and to identify how the disease first appeared in humans. These initial investigations should be followed, not necessarily in any particular order, by the following.
2. Investigations that seek to understand the epidemiology, including how the disease presents clinically in humans, and how infections spread.
3. Studies that investigate potential exposures at the animal–human interface. These would establish which animal species are susceptible and determine whether, where and how the virus might have jumped between those species and humans, or from humans to animals.
4. Studies that identify environmental and ecological influences to determine whether the pathogen is being transmitted by vectors, such as mosquitoes or ticks, or through food or water.
5. Genomic and phylogenetic assessments to establish the pathogen’s closest ancestors and to track its evolution.
6. Investigations that determine whether lapses in biosafety or biosecurity led to an accidental or intentional release of a pathogen into the human population.

This framework has already been applied to the re-emergence of mpox, as well as to the investigation of the origins of SARS-CoV-2. In December 2022, we published recommendations for investigating the origins of mpox (see go.nature.com/3mnmtp).

Animals in the mix

In March this year, we were made aware of new sequencing data released by the Chinese Center for Disease Control and Prevention³, as well as an independent analysis of metagenomics data by international scientists⁴. The analysis assessed wastewater samples, as well as swabs and animal samples that had been collected at the Huanan Seafood Market in December 2019.

The data suggest that a range of animal species, besides those listed in the March 2021 report¹, were present at the market. Although the data do not identify an animal species that was responsible for the first human case of COVID-19, they do point to various animal species as potential intermediate hosts.



Government workers in Côte d'Ivoire assess chicks in a bid to contain an outbreak of H5N1 avian influenza in 2015. Credit: Luc Gnago/Reuters

In our statement this March, SAGO requested that all sequencing data and any preprints that are in review be shared as soon as possible to allow analysis by the scientific community (see go.nature.com/3mddptv). We also requested that researchers in the United States and elsewhere who are using these data collaborate with Chinese researchers.

We will continue to monitor emerging evidence on the origins of SARS-CoV-2 and provide guidance to the WHO on any new data. Continuing from our June 2022 report², we will publish a more comprehensive assessment of current data on the origins of SARS-CoV-2 in the coming months. We will follow up on the recommendations we outlined in June, and will continue to apply the global framework to the investigation of novel pathogens.

For all the work we are doing, however, it is essential that investigations into the origins of novel high-risk pathogens be grounded in science and international collaboration. Such investigations must not be seen as the responsibility of one country, and geopolitics must not be allowed to interfere with or block these.

Throughout the COVID-19 pandemic, attempts to get data out quickly led to the rapid release of preprints, some of which have since been retracted. This has resulted in key stakeholders and the public mistrusting the system. In our view, editors, reviewers and publishers should explore ways of facilitating faster review, to allow rapid sharing of data without erosion of scientific quality.

Platforms such as GISAID (the Global Initiative on Sharing All Influenza Data) have enabled important sequencing data to be shared by scientists, with the aim of also protecting researchers' intellectual property. However, users of these data need to work more closely

with the data owners to maintain trust and ensure that sharing continues in the future. Most importantly, we urge scientists and governments to make available all data, research and reports that can help in the identification of the origins of novel pathogens for all outbreaks, epidemics and global health emergencies.

It is only through these means that investigations can be based on scientific rigour — not on speculations that lead to mistrust.

The authors

Marietjie Venter is chair of the Scientific Advisory Group for the Origins of Novel Pathogens (SAGO) and head of the Zoonotic Arbovirus and Respiratory Virus Research Programme, Centre for Viral Zoonoses, Department of Medical Virology, University of Pretoria, South Africa. See Supplementary information online for a full list of SAGO group members (go.nature.com/43jga3b). e-mail: marietjie.venter@up.ac.za

References

1. World Health Organization. *WHO-Convened Global Study of Origins of SARS-CoV-2: China Part. Joint WHO–China Study* (WHO, 2021).
2. WHO Scientific Advisory Group on the Origins of Novel Pathogens. *Preliminary Report* (WHO, 2022).
3. Liu, W. J. *et al. Nature* <https://doi.org/10.1038/s41586-023-06043-2> (2023).
4. Crits-Christoph, A. *et al.* Preprint at Zenodo <https://doi.org/10.5281/zenodo.7754299> (2023).