

Sharing data to save lives

Journals can and should ensure that they erect no barriers to fast and wide sharing of critical data during major public health emergencies. But funders and scientists must also play a part.

On 1–2 September in Geneva, the World Health Organization (WHO) gathered a group of public health experts, scientists, funders, editors, industry and government representatives for a consultation related to the sharing of data during public health emergencies. Many attendees were directly involved in pathogen genome sequencing and in generating or using clinical, epidemiological and observational data during the 2014 Ebola epidemic in West Africa. Their experiences illustrate that improvements to the norms and process of data sharing in such scenarios are urgently needed.

A notable example was the way in which pathogen genome sequence data was handled. Although two batches of Ebola genome sequences obtained from patient samples were released in April and July 2014, no additional virus genome sequences were released until November that same year. Although restricted availability of patient samples hampered the sequencing of genomes, it is acknowledged that some viral genome sequence data were generated but not publicly released during this interval.

These viral genome sequences would remain useful to investigators seeking to understand the molecular features of virus evolution no matter when they are released. However, their utility to those seeking to trace the chain of transmission (e.g., human-to-human or animal-to-human), to determine whether a pathogen is genetically stable or rapidly mutating, and to track new pathogen variants entering the population, wanes days to weeks after the sequences are generated. Moreover, this ‘real-time’ utility is not restricted to pathogen genome sequence data; clinical trial and observational studies of the efficacy—or lack thereof—of candidate therapies can also enable clinicians to adapt in real-time. Waiting to release these data months after their generation during a public health emergency risks allowing medical workers to repeatedly test therapies that other clinicians had determined were ineffective.

Why are such data not rapidly and widely released in a systematic manner? Many barriers to data sharing were identified during the meeting in Geneva, including patient privacy concerns, government restrictions on the export of patient samples, lack of access to advanced technology in limited resource settings, and lack of personnel and time.

As editors, we wish to address another concern, which a consensus felt was a formidable barrier to early data release: the perception that public release of data will have a negative effect on later scientific publication of the same data. More specifically, many data producers worry that if they publicly release their data prior to submitting a manuscript to a peer-reviewed journal, the editors of the journal may consider the released data no longer novel and reject their manuscript for that reason. In addition, especially when data producers are on the ground working to stem an outbreak and are thus occupied with tasks more important than writing manuscripts, other scientists may analyze the released data, submit the resulting manuscript, and ultimately ‘scoop’ the data-producing lab.

Allaying the first concern is the responsibility of scientific journals. Encouragingly, all editors in attendance emphatically declared that their journals do not wish to hinder real-time sharing of data that could have a clinical impact during public health emergencies. Reflecting this sentiment, the [WHO summary](#) released after the consultation states, “It was unequivocally agreed by representatives from leading biomedical journals that public disclosure of important information of potential relevance to public health emergencies should not be delayed by publication timelines, and that pre-publication disclosure must not and will not prejudice journal publication.” The *Nature* journals agree with this statement; it is consistent with our long-standing policy of not penalizing authors for posting most data on preprint servers. Moving forward, although we already require that all data be available to editors and referees at the time of submission—and freely available to all at the time of publication—when we feel that the data in a submitted manuscript may be useful more broadly to agencies working to stem a public health emergency, we will encourage authors to publicly release the data immediately. We will use editorial discretion in defining major public health emergencies and the relevance of the data in question. To regularly remind readers of these principles, we will ask authors of papers that present data that were released prior to submission or publication to add a sentence to the final published paper explicitly stating this fact.

But these steps address just one part of the problem. Scientists, funders and governments must work to solve the rest. For instance, journals become aware of the existence of data when it is submitted for consideration for publication; yet this could be months after it was initially generated. Only scientists, with the cooperation of governments, can ensure that the data are broadly released as quickly as possible. Scientists are also responsible for ensuring that data users are aware of any limitations of rapidly released potentially preliminary data.

Just as crucially, journals cannot allay concerns about scientists scooping each other. Communities of data producers and users must agree on guidelines for the etiquette of data sharing and data use. Useful precedents might be found in guidelines crafted by other communities that successfully grappled with similar challenges, such as the 2006 [Global Initiative on Sharing Avian Influenza Data](#). Once data producers and users coalesce around a set of guidelines, funders can create policies to ensure that they are followed as often as possible, and that data producers remain incentivized to quickly share their results.

Attendees at the WHO consultation agreed on a list of practical steps that can be taken to remove all identified barriers to data sharing. A list of these steps was submitted to the WHO emergency R&D preparedness blueprint Advisory Group, and it awaits endorsement by the WHO and its partners.

The *Nature* journals’ commitment to enable rather than hinder data sharing in major public health emergencies is effective immediately.