

# Zika-microcephaly paper sparks data-sharing confusion

Virus study used data released online without adequate acknowledgement, researcher complains.

Ewen Callaway

12 February 2016 | Corrected: 12 February 2016



Ernesto Benavides/AFP/Getty Images

Health officials fumigate a cemetery in Lima, Peru, to prevent the spread of mosquito-borne Zika and chikungunya virus.

When researchers in Brazil posted four Zika virus genome sequences in the [online repository GenBank](#) on 26 January, they were complying with a call for scientists to openly release their data during public-health emergencies. By 10 February, the information had been used by Slovenian researchers for their own Zika paper in the *New England Journal of Medicine (NEJM)*<sup>1</sup> — apparently, a textbook example of the power of rapid, open data-sharing.

But the process didn't go entirely smoothly. Oliver Pybus, an evolutionary and infectious-disease biologist at the University of Oxford, UK, who works with the Brazilian group, has complained that the *NEJM* paper did not adequately credit the original data-providers when it only included the GenBank accession number for the data. And Pybus says that he is concerned that this lack of formal recognition could dissuade others from rapidly sharing data during an outbreak.

"The very first big Zika virus paper in the *New England Journal of Medicine* has just created exactly the opposite incentive for groups in Brazil that we want to create. We want them to feel confident they can put their data immediately online without any possible disadvantage to them," Pybus says. The authors of the *NEJM* paper waited to release their own data until their paper was published, he notes.

Tatjana Avšič Županc, a microbiologist at the University of Ljubljana and senior author of the *NEJM* paper, says that her team meant no slight by not contacting Pybus and his colleagues. "If you deposit something in an open domain like GenBank before you publish it, you would expect that people will just use it," she says. And Pybus says that he received an apology from the authors on 11 February, after he contacted them with his concerns. (A representative for the *NEJM* told *Nature* that the scientists should resolve the dispute among themselves.)

But the credit dispute suggests that scientists haven't yet adjusted to the etiquette needed for acknowledging others' public, but not yet formally published, data, researchers say.

**“It might make researchers more aware that sharing data needs to be a reciprocal exercise.”**

### **Microcephaly link**

Županc's team reported a case study of a Slovenian woman who had been living in Brazil and terminated her pregnancy after an ultrasound scan at 29 weeks' gestation revealed that the fetus had microcephaly — an abnormally small head. Zika virus genetic material was discovered in fetal brain tissue and the researchers generated a complete genome sequence.

The team compared their sequence to other Zika sequences in public databases, including the four generated by Pybus's colleague Mario Nunes at the Evandro Chagas Institute in Ananindeua, Brazil, and his team. (In fact, Pybus says, this analysis was not needed to link Zika virus to the case of microcephaly; Avšič Županc says that the analysis was added during the review process at the recommendation of *NEJM* editors).

Pybus and Nunes' team had earlier, on 1 February, posted an online analysis of the data at the website [virological.org](http://virological.org), in which they tracked the importation of Zika to the Americas and its subsequent spread. Pybus thinks that data generated during the Zika virus outbreak is likely to come from a large number of researchers and institutions, which underscores the importance of rapid data sharing.

“I hope it won't discourage people from sharing data pre-publication. It might make researchers more aware that sharing data needs to be a reciprocal exercise,” says Andrew Rambaut, an evolutionary geneticist at the University of Edinburgh, UK. “On the other hand, there is no point in advocating the rapid release of data if you then don't allow people to analyse it. The more eyes that look at it, the more likely that an important finding will be made.”

### **Difficult situation**

Kristian Andersen, an infectious-disease genomicist at the Scripps Research Institute in La Jolla, California, says that the incident highlights “a failure in the system” of using public data that has not yet made it into a publication.

“It's a difficult situation, and our field needs to figure out a way to give better credit to the data producers so data can be shared freely and without limitations,” he says.

Andersen and his colleagues made Ebola virus genome data public during the epidemic in West Africa, and included a note asking scientists who wish to use the data for publication to contact them first. Most scientists who used their data got in touch with them as a result.

Avšič Županc says that there should be clearer standards on how to use another team's unpublished but public data, especially for researchers in fields with different expectations over data sharing.

The incident comes immediately after dozens of funders, government agencies and journals — including the *NEJM* — [released a statement on 10 February](#) supporting open data sharing during public-health emergencies such as the Zika and Ebola epidemics. Journals that signed the statement agreed to make Zika content freely available, and affirmed that early release of data or analysis online will not jeopardize researchers' chances of publication in those journals later on.

*Nature* | doi:10.1038/nature.2016.19367

### **Corrections**

---

**Corrected:** The original version of this article incorrectly stated that the woman who terminated her pregnancy was Brazilian.

### **References**

- 
1. Mlakar, J. *et al.* *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1600651> (2016).