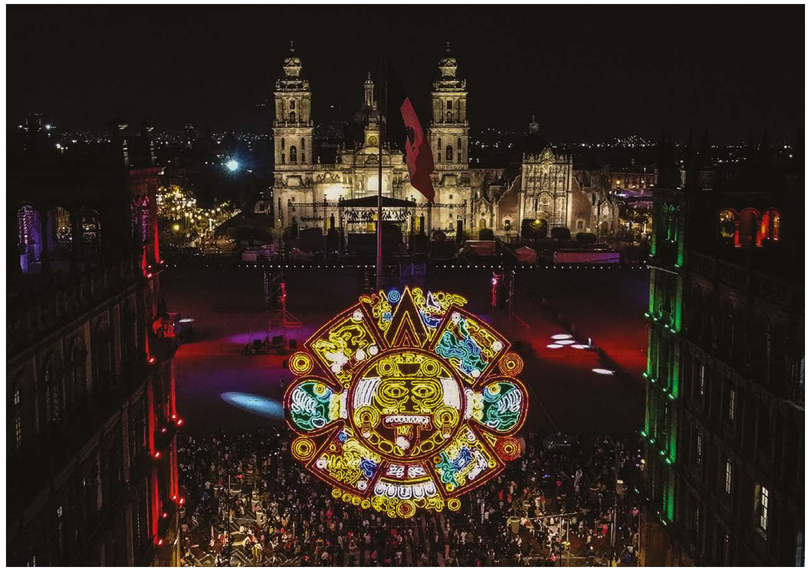
The international journal of science / 19 October 2023

nature

How to share data — not just equally, but equitably

Just as with many natural resources, wealthy countries have been extracting scientific data from poorer nations for centuries. Researchers are changing that.

CRISTOPHER ROGEL BLANQUET/GETTY



Mexico City's people have trusted researchers with their medical histories.

Two decades ago, scientists asked more than 150,000 people living in Mexico City to provide medical data for research. Each participant gave time, blood and details of their medical history.

For the researchers, who were based at the National Autonomous University of Mexico in Mexico City and the University of Oxford, UK, this was an opportunity to study a Latin American population for clues about factors contributing to disease and health. For the participants, it was a chance to contribute to science so that future generations might one day benefit from access to improved health care. Ultimately, the Mexico City Prospective Study was an exercise in trust — scientists were trusted with some of people's most private information because they promised to use it responsibly.

Over the years, the researchers have repaid the communities through studies investigating the effects of tobacco and other risk factors on participants' health[1]. They have used the data to learn about the impact of diabetes on mortality rates[2], and they have found that rare forms of a gene called *GPR75* lower the risk of obesity[3]. And on 11 October, researchers added to the body of knowledge on the population's ancestry[4].

But this project also has broader relevance — it can be seen as a model of trust and of how the power structures of science can be changed to benefit the communities closest to it.

Mexico's population is genetically wealthy. With a complex history of migration and mixing of several populations, the country's diverse genetic resources are valuable to the study of the genetic roots of diseases. Most genetic databases are stocked with data from people with European ancestry. If genomics is to genuinely benefit the global community — and especially under-represented groups — appropriately diverse data sets are needed. These will improve the accuracy of genetic tests, such as those for disease risk, and will make it easier to unearth potential drug targets by finding new genetic links to medical conditions.

At the same time, wealthy nations have, for centuries, been extracting natural resources from low- and middle-income countries and pocketing the proceeds. Scientific data, unfortunately, have been among those resources. Even well-intentioned programmes that emphasize — or require — investigators to share the data from their published studies with the communities from which the information was gathered can reinforce this power structure. Researchers in wealthy nations can access databases more quickly and easily using powerful tools than can those in resource-poor areas; and participating scientists in poorer countries are often not given the credit they are due.

> **"It is an important step towards addressing inequities in data sharing."**

Moreover, if the analyses are designed without the input from the communities from which the data originated, they are less likely to be relevant to that population. This could mean that the promise of serving future generations of Mexicans might go unmet, engendering distrust in scientists and their research aims in the process.

To counter this history, researchers associated with the Mexico City Prospective Study came up with a data-sharing proposal that was expressly designed to benefit scientists in Mexico who want to use the study's data for their own research. The data will be made freely available to researchers in Mexico for up to two years before those in other countries can access them. The authors have also arranged to provide free computing services to those researchers who might want access but who lack the infrastructure to perform analyses on such a massive data set.

There are conditions. Those who wish to benefit will need to ensure that the data are kept securely, and they are not allowed to try to identify the participants. They will also need to seek approval from the original consortium members. Researchers in Mexico who are granted early access to the data can still collaborate with others around the world. It is an important step towards addressing inequities in data sharing that have persisted despite efforts to ensure that the information is shared openly. The study's data-sharing policy builds on previous agreements in which scientists from wealthy countries worked with genomic data from participants in low-income nations.

The solution reached by the Mexico City Prospective

nature

Study is not perfect. Two years of advance access to the data is hardly enough to level the playing field between researchers in Mexico and those in wealthier nations, says Phaik Yeong Cheah, a bioethicist at the Mahidol Oxford Tropical Medicine Research Unit based in Bangkok. To generate genomic data and perform the analyses, the research team partnered with the Regeneron Genetics Center in Tarrytown, New York — an industry-sponsored initiative — and with other pharmaceutical companies. Under the terms of agreement, the initiative will have access to the data from the project alongside scientists in Mexico.

But the spirit of the effort — to specifically consider the ability of researchers to access and analyse the data collected from their communities — is one that more projects could adopt, as they explore ways to counteract the power hierarchies rooted in history that are perpetuated today.

That does not apply only to genetics research. Many fields — including ecology, epidemiology and geology — rely on communities around the world to supply crucial data. It is essential that we explore ways to ensure that those data are used to benefit those who shared them in the first place.

1. Thomson, B. *et al.* *Int. J. Epidemiol.* **50**, 955–964 (2021).
2. Alegre-Díaz, J. *et al.* *N. Engl. J. Med.* **375**, 1961–1971 (2016).
3. Akbari, P. *et al.* *Science* **373**, eabf8683 (2021).
4. Ziyatdinov, A. *et al.* *Nature* https://doi.org/10.1038/s41586-023-06595-3 (2023).

Carl Sagan's audacious search for life on Earth has lessons for today

The test 30 years ago of what remote sensing could tell us about our own planet remains a standard-bearer for careful science.

Early in 1993, a manuscript landed in the *Nature* offices announcing the results of an unusual — even audacious — experiment. The investigators, led by planetary scientist and broadcaster Carl Sagan, had searched for evidence of life on Earth that could be detected from space. The results, published 30 years ago this week, were "strongly suggestive" that the planet did indeed host life. "These observations constitute a control experiment for the search for extraterrestrial life by modern interplanetary spacecraft," the team wrote (C. Sagan *et al.* *Nature* **365**, 715–721; 1993).

The experiment was a master stroke. In 1989, NASA's Galileo spacecraft had launched on a mission to orbit Jupiter, where it was scheduled to arrive in 1995. Sagan and his colleagues wondered whether Galileo would find definitive evidence of life back home if its instruments could be trained on Earth. They persuaded NASA to do just that as the craft flew past the home planet in 1990.

As we describe in an essay on page 451, a big concern for the journal's editors was that the paper did not report a new finding. *Nature* published it because it was a convincing control experiment to test the accuracy and relevance of the methods being used to detect extraterrestrial life. Had the study found less evidence of life than it did, that would have been even more significant — it would have called into question the relevance of the parameters that scientists proposed as evidence of life on other worlds.

> "It is important to test not just what we don't know, but what we think we do know."

Flying visit

The experiment was possible because Galileo had to loop around Earth and Venus on its way to Jupiter, to get a boost from both planets' gravity. It passed 960 kilometres from Earth at its closest point, above the Caribbean Sea.

From the spacecraft's spectrometers, researchers found evidence of oxygen, water vapour, ice and snow, along with carbon dioxide, methane and other greenhouse gases. Its imaging system spotted clouds, oceans, coastlines and rocky surfaces. Although the technology did not have sufficient resolution to be able to detect actual life, it was able to find electromagnetic signals whose amplitude varied in pulses. These amplitude-modulated (AM) waves were used widely at the time to carry radio and television broadcasts, and were of a type not known to occur naturally. "Of all Galileo science measurements, these signals provide the only indication of intelligent, technological life on Earth," the team wrote. This was a delicious twist, because Sagan was constantly on radio and television, one of the most recognized science broadcasters of his generation.

The study, now commonly taught, has stood the test of time and contributed to further thinking on frameworks for reporting evidence of life on other planets. Since the early 1990s, astronomers have discovered more than 5,500 planets orbiting stars outside the Solar System. Moreover, a cascade of discoveries is expected in data from NASA's powerful James Webb Space Telescope (JWST), which is uniquely well equipped to study exoplanet atmospheres.

Three decades on, Sagan's classic experiment has three important lessons for researchers and science publishers. The first is that it is important to test not just what we don't know, but what we think we do know. The second is a reminder to those of us who publish science that control experiments — like replication studies — are as important as research that describes new results.

Last but not least is the lesson implicit in the great care the team took when reporting the findings, including the detection of chemical signatures such as the presence of water or greenhouse gases. It would have been easy, given what was known about life on Earth, to assume that the first piece of evidence clinched the question. Instead, the researchers built up a nuanced conclusion, bringing together all the evidence available to them. Their approach demonstrates why the search for extraterrestrial life will always be one of the hardest problems in science.