

New data-sharing requirements from the National Institutes of Health are a big step toward more open science

March 24 2022, by Stephen Jacobs



The National Institutes of Health has had data-sharing guidelines in place for years, but the new rules are by far the most comprehensive. Credit: [NIH](#)

Starting on Jan. 25, 2023, many of the 2,500 institutions and 300,000 researchers that the U.S. [National Institutes of Health](#) supports will need to provide a formal, detailed plan for publicly sharing the data generated by their research. For many in the scientific community, this new NIH [Data Management and Sharing Policy](#) sounds like a no-brainer.

The incredibly quick development of rapid tests and vaccines for COVID-19 demonstrate the success that can follow the open sharing of [data](#) within the [research community](#). The importance and impact of that data even drove a White House [Executive Order](#) mandating that "the heads of all executive departments and agencies" share "COVID-19-related data" publicly last year.

I am the Director of the Rochester Institute of Technology's Open Programs Office. At [Open@RIT](#), my colleagues and I work with faculty and researchers to help them openly share their research and data in a manner that provides others the rights to access, reuse and redistribute that work with as few barriers or restrictions as possible. In the sciences, these practices are often referred to as [open data](#) and [open science](#).

The journal *Nature* has called the impact of the NIH's new data management policy "seismic," saying that it [could potentially create a "global standard" for data sharing](#). This type of [data sharing](#) is likely to produce many benefits to [science](#), but there also are some concerns over how researchers will meet the new requirements.

What to share and how to share it

The NIH's new policy around data sharing replaces [a mandate from 2003](#). Even so, for some scientists, the new policy will be a big change. Dr. Francis S. Collins, then Director of the NIH, said in the 2020 statement [announcing the coming policy changes](#) that the goal is to "shift the culture of research" so that data sharing is the norm, rather than the

exception."

Specifically, the policy requires two things. First, that researchers share all the [scientific data](#) that other teams would need in order to "validate and replicate" the original research findings. And second, that researchers include a two-page [data management plan](#) as part of their application for any NIH funding.

So what exactly is a data management plan? Take an imaginary study on heat waves and heatstroke, for example. All good researchers would collect measurements of temperature, humidity, time of year, weather maps, the health attributes of the participants and a lot of other data.

Starting next year, research teams will need to have determined what reliable data they will use, how the data will be stored, when others would be able to get access to it, whether or not special software would be needed to read the data, where to find that software and many other details—all before the research even begins so that these things can be included in the proposal's data management plan.

Additionally, researchers applying for NIH funding will need to ensure that their data is available and stored in a way that persists long after the initial project is over.

The NIH has stated that it will support—with additional funding—the costs related to the [collection, sharing](#) and [storing of data](#).

Sharing data promotes open science

The NIH's case for the new policy is that it will be "[good for science](#)" because it maximizes availability of data for other researchers, addresses problems of reproducibility, will lead to better protection and use of data and increase transparency to ensure public trust and accountability.

The first big change in the new policy—to specifically share the data needed to validate and replicate—seems aimed at the [proliferation of research that can't be reproduced](#). Arguably, by ensuring that all of the relevant data from a given experiment is available, the scientific world would be better able to evaluate and validate through replication the quality of research much more easily.

I strongly believe that requiring data-sharing and management plans addresses a big challenge of open science: being able to [quickly find the right data, as well as access, and apply it](#). The NIH says, and I agree, that the requirement for data management plans will help make the use of open data faster and more efficient. From the Human Genome Project in the 1990s to the recent, rapid development of tests and vaccines for COVID-19, the benefits of greater openness in science have been borne out.

Will the new requirements be a burden?

At its core, the goal of the new policy is to make science more open and to fight bad science. But as beneficial as the new policy is likely to be, it's not without costs and shortfalls.

First, replicating a study—even one where the data is already available—still consumes expensive human, computing and material resources. The system of science doesn't reward the researchers who reproduce an experiment's results as highly as the ones who originate it. I believe the [new policy](#) will improve some aspects of replication, but will only address a few links in the overall chain.

Second are concerns about the increased workload and financial challenges involved in meeting the requirements. Many scientists aren't used to preparing a detailed plan of what they will collect and how they will share it as a part of asking for funding. This means they may need

training for themselves or the support of trained staff to do so.

Part of a global trend toward open science

The NIH isn't the only federal agency pursuing more open data and science. In 2013, the Obama administration mandated that all agencies with a budget of \$100 million or more must provide open access to their publications and data. The National Science Foundation [published their first open data policy](#) two years earlier. Many European Union members are crafting national policies on open science—most notably France, which has already [published it's second](#).

The cultural shift in science that NIH Director Collins mentioned in 2020 has been happening—but for many, like me, who support these efforts, the progress has been painfully slow. I hope that the new NIH open data [policy](#) will help this movement gain momentum.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: New data-sharing requirements from the National Institutes of Health are a big step toward more open science (2022, March 24) retrieved 24 April 2024 from <https://medicalxpress.com/news/2022-03-data-sharing-requirements-national-health-big.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--