

# Clarifying access to data

To facilitate access to and improve the discoverability of the data in our papers, *Nature Medicine* is making the data availability statement in our papers more prominent and its language more transparent.

Sharing of research data is a condition for publication in *Nature Research journals*, including *Nature Medicine*, and for receiving grants from many funding agencies, including the US National Science Foundation, the US National Institutes of Health, the Wellcome Trust, Research Councils UK and the Bill & Melinda Gates Foundation. Beyond this, data sharing is critical for reproducibility, the oil that keeps the wheels of science turning. The ability of researchers to reproduce the findings of scientific studies not only strengthens the confidence of the public and the scientific community in reported results, but also helps ensure that science as a whole progresses and benefits society.

Over the past decade, there have been increasing calls to improve the reproducibility of experimental findings through greater data transparency and accessibility. *Nature Research journals* have responded to this need and have taken steps to ensure greater [transparency in data reporting](#), such as by asking authors to provide details about elements of experimental and analytical design that are frequently under-reported. These reporting summaries are then published alongside the final paper. We have also taken steps to improve the way data are represented in our papers: for instance, by requiring authors to present data in a format that shows data distribution and to provide full western blot images as part of the supplementary material, as well as by encouraging authors to include raw data tables with the accepted paper and to publish Data Descriptors of their datasets in our sister journal *Scientific Data*.

But data transparency also means ensuring access to the data behind published papers to the scientific community. Thus, consistent with our drive to support reproducibility and robustness in research, in 2016 we began requiring our authors to include a data availability statement (DAS) as part of the online methods section of their published

papers. This statement consolidates information on the availability of data associated with the manuscript and describes whether the data are in repositories, available on request, included with supplementary information or in figure source data files. The DAS has the purpose of offering a clear path for the community to access the raw data underlying the results reported. Now, to further enhance data discoverability, *Nature Research journals* are increasing the prominence of the DAS by making it an independent section of the paper that is available without restriction to all readers.

Taking this initiative one step further, we at *Nature Medicine* will be working with authors to ensure that the DAS is detailed and transparent, provides essential information on where the data are and how they can be accessed, and discloses any restrictions that may exist with regard to making the data available to other investigators. Historically, we have accepted a minimal DAS that simply affirms that the data supporting the findings of the study are available from the corresponding author upon reasonable request. However, our experience has been that the definition of a ‘reasonable request’ is often subjective and does not adequately reflect the complexities associated with the types of studies that we publish.

For instance, data associated with research involving human participants are subject to government regulations in the US (in particular, the Health Insurance Portability and Accountability Act) and the European Union (under the General Data Protection Regulation). To protect patients’ privacy, measures must be taken in the collection, storage and reporting of patient information, including not just patients’ consent for the use of their data for research but also deidentification of patients from their underlying data. Access to patient data—whether they are deposited in repositories with controlled access, such as

the database of Genotypes and Phenotypes (dbGaP) or the European Genome-phenome Archive (EGA), or provided upon direct request to investigators—requires additional steps of internal board review and approval to ensure that the purpose of a request is legitimate and bound to the terms consented by patients. In other cases, accessibility may be affected by concerns about commercial use of the data and sharing of proprietary technology or datasets. Therefore, we appreciate that not all data are sharable without restrictions.

While there are times when the data supporting the results in our pages cannot be freely shared, we wish to make these restrictions more apparent to our readers and we think that authors should be more transparent with regard to any restrictions on data availability that may exist. The DAS offers the proper venue for such disclosure. Authors should provide a summary of the process for requesting access to their data, identify the entity at their institution or company that will review such requests, detail how requests will be measured against privacy and/or intellectual property concerns, and describe by what means data or material will be shared (for example, whether a material transfer agreement will be required).

The late US Supreme Court associate justice Louis D. Brandeis once remarked that “sunlight is said to be the best of disinfectants” with regard to maintaining the transparency of government policy and thus ensuring good governance. In a similar vein, we expect authors to be transparent about their data so that science can benefit and progress that much further and faster. We hope that the changes we are making to the DAS will raise reader awareness of this information and, ultimately, that the DAS will facilitate access and enhance the discoverability of the source data for our papers. □

Published online: 7 September 2018  
<https://doi.org/10.1038/s41591-018-0188-2>