



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Open Science at PLOS Pathogens

**Citation for published version:**

Cadwallader, L, Haldar, K, Kirk, R, Mabbott, NA & Malim, MH 2023, 'Open Science at PLOS Pathogens', *PLoS Pathogens*, vol. 19, no. 11, e1011776, pp. 1-6. <https://doi.org/10.1371/journal.ppat.1011776>

**Digital Object Identifier (DOI):**

[10.1371/journal.ppat.1011776](https://doi.org/10.1371/journal.ppat.1011776)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

PLoS Pathogens

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## EDITORIAL

Open Science at *PLOS Pathogens*

Lauren Cadwallader<sup>1\*</sup>, Kasturi Haldar<sup>2,3</sup>, Rebecca Kirk<sup>1</sup>, Neil A. Mabbott<sup>4</sup>, Michael H. Malim<sup>2,5</sup>

**1** PLOS, Cambridge, United Kingdom, **2** Editor-in-Chief, *PLOS Pathogens*, PLOS, United Kingdom, **3** University of Notre Dame, Notre Dame, Indiana, United States of America, **4** The Roslin Institute & Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, United Kingdom, **5** King's College London, London, United Kingdom

\* [lcadwallader@plos.org](mailto:lcadwallader@plos.org)

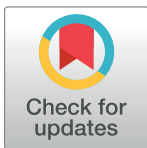
The pathogens research community has driven, and benefited from, Open Science. As a leader in the field, *PLOS Pathogens* strives to advance community-rooted adoption of practices that enable transparency, rapid communication, reproducibility, and trust in research, and help to transform research globally. A key vision of *PLOS Pathogens* is to facilitate widespread adoption of Open Science to accelerate and improve research that is rigorous and trustworthy and to ensure its meaningful impact on lives.

### The importance of Open Science

Open Science—the public, unrestricted sharing of all components of research—can improve science in many ways and has benefits for a wide range of stakeholders. Some elements of Open Science, such as Open Access to publications, advance equity to ensure research findings can be accessed by everyone with a vested interest in the science. Other elements, such as resource and data sharing, enable researchers to reuse and build on existing science. What all elements have in common is the ability to increase trust, reproducibility, and transparency in science (Fig 1). These attributes are essential not only for the scientific community but also for policymakers, funders, practitioners, and the public. Open Science is rooted in the principles of fairness, justice, and equity and thus enables and empowers global and diverse communities to participate in and contribute to more efficient science [1].

Open Science has long been practiced by the pathogens research community, given its importance in the context of the challenges we seek to address, and the collaborative nature and potential impacts our research can have on the lives of people across the globe. *PLOS Pathogens* was founded with a philosophy based on open, high-quality, rigorous science, initially focusing on Open Access publication and diversifying the ways scientists communicate with the community and public (for example, see [2]) and more recently experimenting with ways to enhance data sharing [3]. We have highlighted the importance of Open Science to the community and discipline since the inception of the journal [4] and, over the years, advocated for openness via the research articles published in the journal (for example, [5,6]). *PLOS Pathogens'* position on Open Science also means it is already compliant with established and emerging funder policies around Open Science, such as Plan S [7] and the OSTP memo [8], that will bring United States-funded research in line with European Union and United Kingdom funders and further the adoption of Open Science.

Beyond *PLOS Pathogens*, we have recently seen the critical importance of Open Science practices including the sharing of preprints and the genomic sequencing data of SARS-CoV-2 to drive rapid innovation in research during the global response to the COVID-19 pandemic [9,10]. The AlphaFold Protein Structure Database [11] has over 350,000 protein models



### OPEN ACCESS

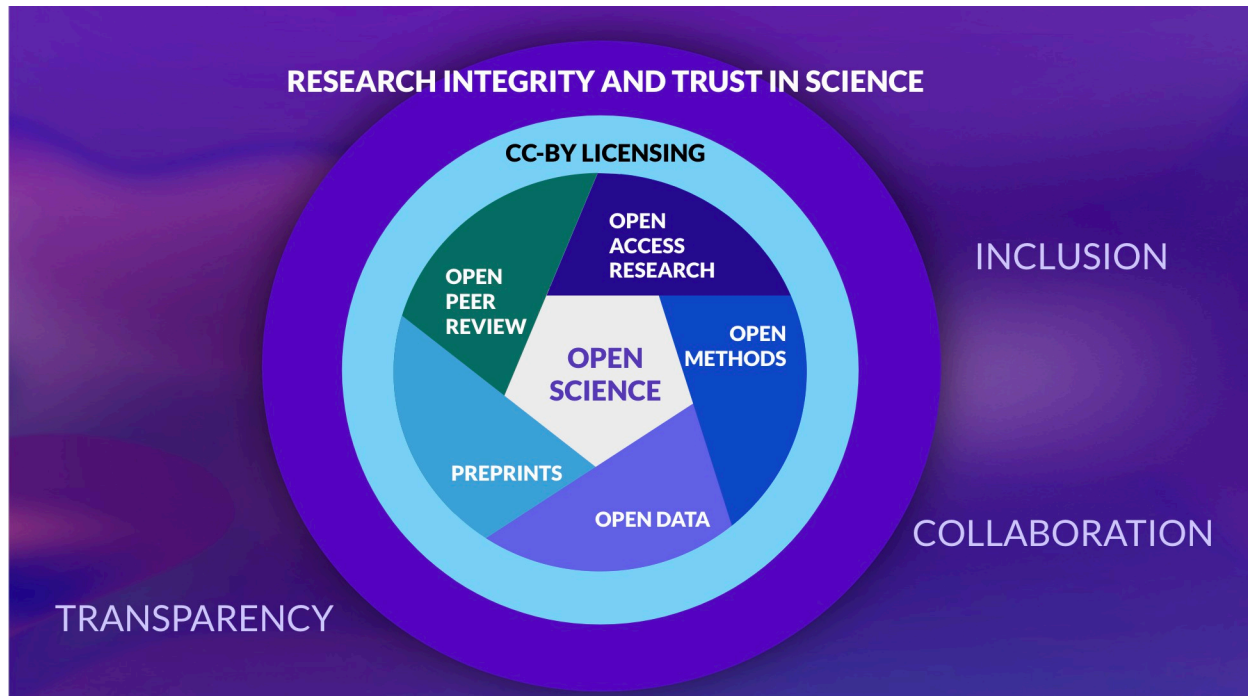
**Citation:** Cadwallader L, Haldar K, Kirk R, Mabbott NA, Malim MH (2023) Open Science at *PLOS Pathogens*. *PLoS Pathog* 19(11): e1011776. <https://doi.org/10.1371/journal.ppat.1011776>

**Published:** November 30, 2023

**Copyright:** © 2023 Cadwallader et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.



**Fig 1. The interrelated parts of Open Science.**

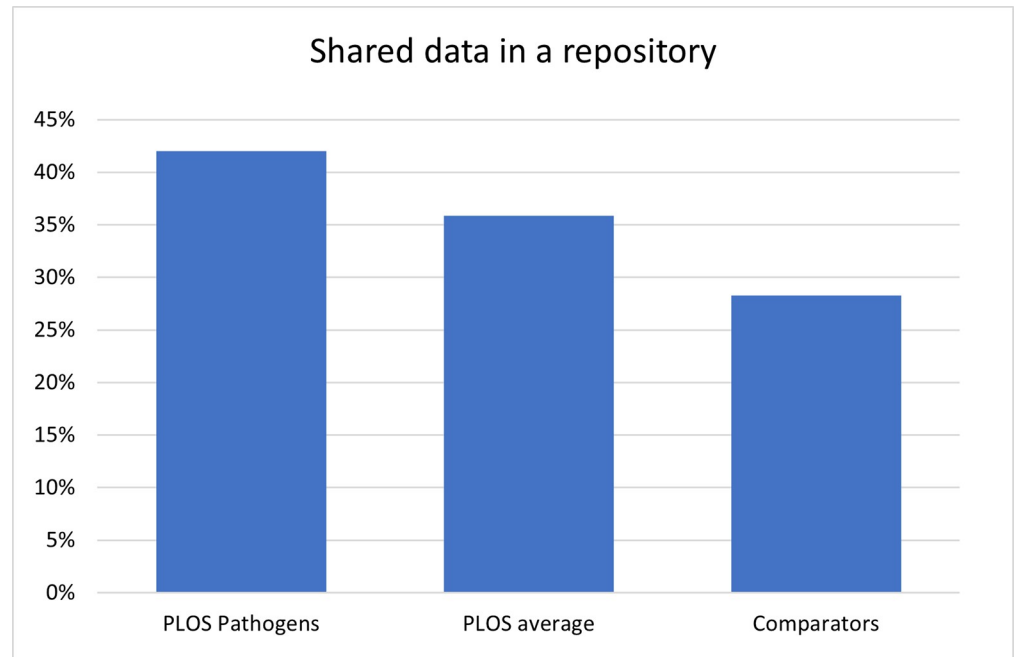
<https://doi.org/10.1371/journal.ppat.1011776.g001>

available, as well as the tools to view and interrogate the structure, that is available with an Open Access license. This enables and accelerates research into protein structures, for example, in response to the emergence of SARS-CoV-2, as well as in other areas related to pathogens research.

Funders and major journals require that data, such as proteomics and transcriptomics data, as well as new computer code are deposited in public repositories. The vast array and public availability of these primary datasets, as well as the stabilization of analysis platforms and methodology, permits retrospective meta-analysis of the wealth of useful data produced in different laboratories. For example, meta-analyses of many previously published and publicly available datasets from diverse tissues and cell lineages can be used to identify clusters of genes that are correlated in their expression across large datasets, which, in turn, can give clear insights into the transcriptional networks underlying specific tissue or cells (hematopoietic, mesenchymal lineages, etc.) or cellular functions (phagocytosis, cell-cycle, etc.). The resulting dataset can, in turn, be shared (for example, [12]) and used to look at numerous research questions. For example, by identifying novel intervention strategies to reverse the effects of immunosenescence on antigen-specific immunity in the skin of elderly humans *in vivo* [13], to help understand the transcriptional landscapes of thyroid cancer [14] and hepatocellular carcinoma [15], and to compare the cellular composition of whole blood samples from pediatric septic shock patients based on analysis of mRNA expression profiles [16].

### Cooperative and cocreative Open Science

The pathogens community has, and continues to, demonstrate its commitment to cooperative Open Science to improve and enhance science. One example would be the many community-curated virus databases used in pathogens research (see [17] for a review). A number of organizations, including the NIH HIV Reagent Program [18], the Biodefense & Emerging Infections



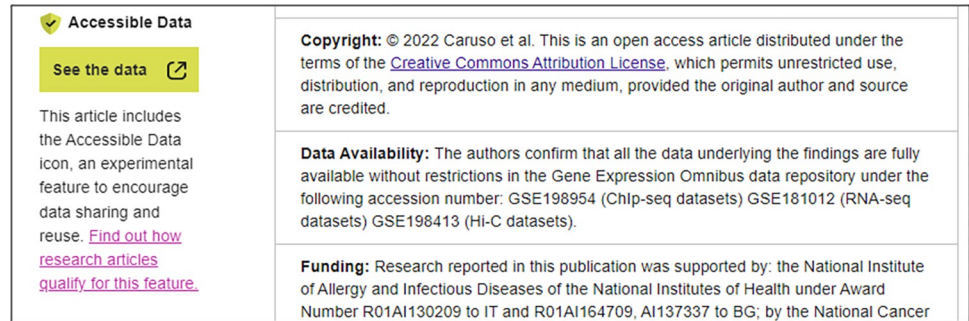
**Fig 2. The percentage of shared data that is available from a repository for *PLOS Pathogens*, the average across PLOS journals and in comparator articles for research articles published between 1 Jan 2019 and 31 Mar 2023. Data from [27].**

<https://doi.org/10.1371/journal.ppat.1011776.g002>

Resources [19], the International Reagent Resource [20], the National Institute for Biological Standards and Control [21], and International Human Papillomavirus Reference Center [22], all freely provide reagents and standards that are vital to the pursuit of pathogen-focused research and have been provided by researchers in the field. Without this open sharing, research would be limited as it is reliant on the availability of high-quality and authenticated research materials. Similarly, Addgene [23], a nonprofit plasmid repository, facilitates the storage and distribution of plasmids across the globe that have been deposited by researchers for use by others. However, fees borne by the reuser, such as shipping costs, still act as a barrier to truly equitable Open Science. There is an opportunity here for the pathogens community to demonstrate their commitment further by creating pathways for researchers who still face barriers relating to the cost of research.

Cocreation is central to PLOS' approach to Open Science. Open Science solutions will be most effective when they are initiated by the community and work with established practices and community norms. *PLOS Pathogens* embraced this approach with the adoption of the PLOS-wide mandatory data sharing policy in 2014, created after extensive consultation with the academic community [24]. Since 2014, the rate of data sharing, and data shared following best practice, has been increasing [25].

We have worked with the community to offer experimental solutions we think have potential and to learn about why particular solutions do (or do not) work. PLOS can now monitor the effect of Open Science practices of authors, thanks to the Open Science Indicators initiative [26]. For example, we know that 48% of *PLOS Pathogens* authors sharing data use a repository compared to only 28% of authors in the broader scholarly community (see Fig 2) [27], and we can use this data to drill down further and understand which repositories our authors use. This has helped shape the extension of one of our new experimental features—the Accessible Data icon (see Fig 3)—to include the most commonly used repositories at PLOS, including



**Fig 3. A snippet of a *PLOS Pathogens* article (<https://doi.org/10.1371/journal.ppat.1010400>), which now features the Accessible Data icon that takes readers straight to the data cited in the Data Availability statement.**

<https://doi.org/10.1371/journal.ppat.1011776.g003>

Dryad, Figshare, Open Science Framework (OSF), Github, Zenodo, Gene Expression Omnibus, Sequence Read Archive, BioProject, and Demographic and Health Surveys [28]. Following the extension of the Accessible Data icon to include more repositories, the number of *PLOS Pathogens* articles that feature the Accessible Data icon has increased by almost 20 times. Results from the experiment have shown that datasets connected to articles carrying the Accessible Data feature have had significantly more engagement since the addition of the feature. In addition, qualitative results from the experiments support the potential of the icon to normalise data sharing and influence author behaviour [28].

Open, high-quality, peer-reviewed science is, and has always been, core to the mission of *PLOS Pathogens*, and we hope the journal can continue to work with the community to advance Open Science for years to come.

## Acknowledgments

We would like to thank the following *PLOS Pathogens* Editorial Board Members for comments on an earlier draft of this Editorial: Sonja Best, Jeffrey Dvorin, Alexander Gorbalenya, and Alison McBride. We would also like to thank Julia Squarr and Lindsay Morton from *PLOS* for their input.

## Author Contributions

**Conceptualization:** Lauren Cadwallader, Kasturi Haldar, Rebecca Kirk, Michael H. Malim.

**Writing – original draft:** Lauren Cadwallader.

**Writing – review & editing:** Lauren Cadwallader, Kasturi Haldar, Rebecca Kirk, Neil A. Mabbott, Michael H. Malim.

## References

1. Open Science by Design: Realizing a Vision for 21st Century Research. Washington, DC.: National Academies Press; 2018. <https://doi.org/10.17226/25116>
2. Madhani HD, Haldar K, McFadden G. “Pearls”: A New Type of Open-Access Educational Resource. *PLoS Pathog.* 2009; 5(6):e1000499. <https://doi.org/10.1371/journal.ppat.1000499>
3. Cadwallader L, Haldar K, Malim MH. Increasing best practice data sharing at *PLOS Pathogens*. *PLoS Pathog.* 2021; 17(11):e1010021. <https://doi.org/10.1371/journal.ppat.1010021> PMID: 34762715
4. Coloma J, Harris E. Open-Access Science: A Necessity for Global Public Health. *PLoS Pathog.* 2005; 1(2):e21. <https://doi.org/10.1371/journal.ppat.0010021> PMID: 16254599



5. Van Voorhis WC, Adams JH, Adelfio R, Ahyong V, Akabas MH, Alano P, et al. Open Source Drug Discovery with the Malaria Box Compound Collection for Neglected Diseases and Beyond. *PLoS Pathog.* 2016; 12(7):e1005763. <https://doi.org/10.1371/journal.ppat.1005763> PMID: 27467575
6. Baker D, van den Beek M, Blankenberg D, Bouvier D, Chilton J, Coraor N, et al. No more business as usual: Agile and effective responses to emerging pathogen threats require open data and open analytics. *PLoS Pathog.* 2020; 16(8):e1008643. <https://doi.org/10.1371/journal.ppat.1008643> PMID: 32790776
7. European Science Foundation. Plan S Principles. [cited 2023 Jul 13]. Available from: [https://www.coalition-s.org/plan\\_s\\_principles/](https://www.coalition-s.org/plan_s_principles/)
8. White House Office of Science and Technology Policy (OSTP). Desirable Characteristics of Data Repositories for Federally Funded Research. Executive Office of the President of the United States; 2022 May. <https://doi.org/10.5479/10088/113528>
9. Morton L. Reflecting and expanding on PLOS Pathogens' response to COVID-19. Speaking of Medicine and Health [Internet]. 2021 Jun 14 [cited 2023 Jul 13]. Available from: <https://speakingofmedicine.plos.org/2021/06/14/pathogens-covid-update/>
10. Corbett KS, Edwards DK, Leist SR, Abiona OM, Boyoglu-Barnum S, Gillespie RA, et al. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature.* 2020; 586:567–571. <https://doi.org/10.1038/s41586-020-2622-0> PMID: 32756549
11. European Molecular Biology Laboratory. AlphaFold Protein Structure Database. [cited 2023 Oct 24]. Available from: <https://alphafold.ebi.ac.uk/>
12. Mabbott NA, Baillie J, Brown H, Freeman TC, Hume DA. An expression atlas of human primary cells: inference of gene function from coexpression networks. *BMC Genomics.* 2013; 14:632. <https://doi.org/10.1186/1471-2164-14-632> PMID: 24053356
13. Chambers ES, Vukmanovic-Stejic M, Shih BB, Trahair H, Subramanian P, Devine OP, et al. Recruitment of inflammatory monocytes by senescent fibroblasts inhibits antigen-specific tissue immunity during human aging. *Nat Aging.* 2021; 1:101–113. <https://doi.org/10.1038/s43587-020-00010-6> PMID: 37118005
14. Hwang E, Doolittle WKL, Zhu YJ, Zhu X, Zhao L, Yu Y, et al. Thyroid hormone receptor  $\alpha 1$ : a novel regulator of thyroid cancer cell differentiation. *Oncogene.* 2023; 42:3075–3086. <https://doi.org/10.1038/s41388-023-02815-2> PMID: 37634007
15. Li S, Du H, Gan D, Li X, Zao X, Ye Y. Integrated analysis of single-cell and bulk RNA-sequencing reveals tumor heterogeneity and a signature based on NK cell marker genes for predicting prognosis in hepatocellular carcinoma. *Front Pharmacol.* 2023; 14:1200114. <https://doi.org/10.3389/fphar.2023.1200114> PMID: 37397471
16. Yang JO, Zinter MS, Pellegrini M, Wong MY, Gala K, Markovic D, et al. Whole Blood Transcriptomics Identifies Subclasses of Pediatric Septic Shock. *Res Sq.* 2023 Aug. <https://doi.org/10.21203/rs.3.rs-3267057/v1> PMID: 37693502
17. Ritsch M, Cassman NA, Saghaei S, Marz M. Navigating the Landscape: A Comprehensive Review of Current Virus Databases. *Viruses.* 2023; 15:1834. <https://doi.org/10.3390/v15091834> PMID: 37766241
18. National Institute of Health. NIH AIDS Reagent Program. [cited 2023 Oct 23]. Available from: <https://www.niaid.nih.gov/research/nih-aids-reagent-program>
19. National Institute of Health. BEI Resources Repository. [cited 2023 Oct 24]. Available from: <https://www.niaid.nih.gov/research/bei-resources-repository>
20. Centers for Disease Control and Prevention. International Reagent Resource (IRR). [cited 2023 Oct 23]. Available from: <https://www.internationalreagentresource.org/>
21. National Institute for Biological Standards and Control (NIBSC). National Institute for Biological Standards and Control (NIBSC). [cited 2023 Oct 18]. Available from: <https://www.nibsc.org/>
22. International Human Papillomavirus Reference Center. International Human Papillomavirus Reference Center. [cited 2023 Oct 24]. Available from: <https://www.hpvcntr.se/>
23. Addgene. Addgene repository. [cited 2023 Oct 16]. Available from: <https://www.addgene.org/>
24. Bloom T, Ganley E, Winker M. Data Access for the Open Access Literature: PLOS's Data Policy. *PLoS Biol.* 2014; 12(2):e1001797. <https://doi.org/10.1371/journal.pbio.1001797>
25. Cadwallader L, Morton L, Hrynaszkiewicz I. A new Open Science Indicators dataset is here! The Official PLOS Blog [Internet]. 2023 Oct 3 [cited 2023 Oct 24]. Available from: <https://theplosblog.plos.org/2023/10/open-science-indicators-q2-2023/>
26. Cadwallader L, Morton L, Hrynaszkiewicz I. A first look at Open Science Indicators for articles published in 2023. The Official PLOS Blog [Internet]. 2023 Jun 26 [cited 2023 Jul 5]. Available from: <https://theplosblog.plos.org/2023/06/open-science-indicators-update-q1-2023/>

27. Public Library of Science. PLOS Open Science Indicators (version 3). Figshare. 2022. <https://doi.org/10.6084/m9.figshare.21687686.v3>
28. Hrynaszkiewicz I. Extending Accessible Data to more articles, repositories, and outputs. The Official PLOS Blog [Internet]. 2023 Jul 6 [cited 2023 Jul 14]. Available from: <https://theplosblog.plos.org/2023/07/accessible-data/>.