

Sharing study materials in health and medical research

Nicholas J. DeVito^{*1}, Caroline Morton¹, Aidan G. Cashin^{2,3}, Georgia C. Richards⁴, Hopin Lee^{6,7}

1. Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford UK.
2. Centre for Pain IMPACT, Neuroscience Research Australia, Randwick, New South Wales, Australia
3. School of Health Sciences, University of New South Wales, Sydney, New South Wales, Australia
4. Centre for Evidence Based Medicine, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
5. Centre for Statistics in Medicine & Rehabilitation Research in Oxford, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Oxford, UK
6. School of Medicine and Public Health, The University of Newcastle, Callaghan, New South Wales, Australia

Word count: 2800

* Corresponding Author

Nicholas J. DeVito

nicholas.devito@phc.ox.ac.uk

Abstract

Making study materials available allows for a more comprehensive understanding of the scientific literature. Sharing can take many forms and include a wide variety of outputs including code and data. Biomedical research can benefit from increased transparency but faces unique challenges for sharing, for instance confidentiality concerns around participant's medical data. Both general and specialised repositories exist to aid in sharing most study materials. Sharing may also require skills and resources to ensure that it is done safely and effectively. Educating researchers on how to best share their materials, and properly rewarding these practices, requires action from a variety of stakeholders including journals, funders, and research institutions.

Key Messages

What is already known on this topic: Sharing study materials can have numerous advantages however in practice sharing is inconsistent and understudied.

What this study adds: This paper introduces the basic arguments in favour of increased sharing, provides guidance on how to safely and effectively share, while also examining barriers that may impact sharing.

How this study might affect research, practice or policy: Increased awareness and education about sharing practices allows researchers to better consider how, where, and when they might aim to share their research materials.

Introduction

Clear and accurate descriptions of methods are essential to the scientific endeavour. However, studies can range broadly in their complexity and the subsequent level of detail required to fully understand the methods. Relying solely on space-limited narrative methods sections, which some journals relegate to the end of the manuscript, can leave gaps for readers that make it difficult to assess what occurred in practice, and if there were any potential issues in the execution of a given study. This leads to difficulties in properly assessing a study's merits, reproducing its findings, and confidently building upon its conclusions.[2–4].

Large scale investigations of transparency practices in the biomedical literature, including sharing, have shown improvements over time but with room for further growth.[38] Sharing of data, code, and protocols lags “significantly behind” other areas of transparency practice including disclosures of conflicts of interest and funding in the literature.[39] Investigations of the data sharing standards in health and medical journals have also shown clear gaps [40,41] and there might be a mismatch between the willingness to share and actual sharing rates[42,43].

Examinations of the data sharing policies of commercial and non-commercial research sponsors, as well as journals, show high variability in the extent that data sharing is encouraged, supported, or required.[44–49]. Moreover, problems with reproducibility have been identified in clinical trials [50], observational research [51], and systematic reviews [52]. This matches the low overall computational reproducibility seen in the broader literature.[53]

To ensure a sufficient level of detail is made available, many have called for the sharing of study materials to become more routine, expected, and rewarded throughout science.[5] Starting in 2023, plans for sharing and management of data will be a technical aspect of grant assessments at the US National Institutes of Health, the largest health funder in the world, while other major funders, like the UK National Institute for Health and Care Research promote the sharing of funded research outputs.[6,7] What exactly should be shared, and how to accomplish this increased transparency, will vary considerably both across, and within, disciplines.[8]

This article addresses sharing in the context of general health and medical research. We focus on the sharing of study data, analysis code, and touch upon other study-specific materials; however, the general advice will cover most materials regardless of study-level idiosyncrasies.

What does ‘sharing’ mean?

When discussing ‘sharing’ in the context of academic investigations, we mean making relevant materials, key to understanding the conduct and analysis of a study, available to interested parties to the extent ethically and legally possible (**Box 1**). Fundamental to sharing of study materials is that they are made available as part of the timely and complete dissemination of results. While we focus on sharing as it relates to data and materials either before or alongside results in this piece, biases in the availability of results will undermine any efforts for more open and transparent science. For many, sharing in the context of research means making their raw data available. Sharing data allows for external review and reuse to support future research.[9] In 2018, the International Committee of Medical Journal Editors (ICMJE) began requiring that articles reporting clinical trials in member journals must contain a data sharing statement.[10] Notably, this is not a requirement to share data, but rather an attempt to be transparent about the extent or conditions around data sharing. However, there is no guarantee these statements will be implemented, honoured, or to lead to increased sharing.[11,12]

Box 1: Summary of Sharing Practices

What: Sharing data, code, and other study materials like protocols, documentation, and guides.

Why: Increased transparency of research allows critical appraisal, reproducibility, and secondary analysis that increases confidence in results and extends its value.

When: Sharing can occur at any time throughout a study with some expected at the start (e.g., protocols), some potentially during a study (e.g., preliminary data, guides) and some in a timely manner after completion of the study (e.g., final data and analysis code).

Who: Investigators usually bear the final responsibility for their data but institutions, funders, and journals, including both editors and peer reviewers, can aid in ensuring sharing occurs efficiently, effectively, and safely.

How: Sharing happens in a number of ways specific to a given study but various platforms, technologies, and repositories can handle a wide range of study materials and aid in making them discoverable and appropriately accessible.

While data sharing is important, it is only one aspect of the study methods that can be made available.[13] A wide array of documentation, data collection sheets and instruments, analysis plans and code, and various other forms of information can be shared to provide insights into how a study was conducted (**Box 2**). The Transparency and Openness Promotion (TOP) Guidelines grade journal policies across eight standards, including the sharing of data, analytic code, materials, and documentation (i.e., protocols and analysis plans) and each area is assessed across four levels ranging from “not implemented” to having an active requirement and verification of sharing.[14]

Box 2: A Case Study in Sharing Materials in Health and Medicine

Eklund and colleagues published an analysis of common software packages for functional MRA (fMRI) research. By running tests on an fMRI dataset, they found issues with the implementation of these programs and concluded that there was a need to validate “the statistical methods being used in the field of neuroimaging.”

The study used an open fMRI dataset* and the group shared all their processing scripts on GitHub.** This meant that when the article’s findings generated attention within the field, others were able to use these scripts and data to contextualise, replicate, verify, and extend the original claims. Furthermore, because some of the fMRI statistical software is maintained

as an open source software project, the teams were able to find and fix bugs that could impact future analyses.

Eklund and colleagues end their response to comments on their original piece by noting that “together, these examples show the importance of data sharing, open-source software, code sharing, and reproducibility.”[15]

*<https://www.openfmri.org/faq/>

**<https://github.com/wanderine/ParametricMultisubjectfMRI>

Why Should Study Materials be Shared?

It is common practice in manuscripts with data sharing statements to make even public, non-sensitive data, or other study materials, available only by “reasonable requests” to an author which can pose barriers between potential users and study materials [21][22]. Email requests for access can go to defunct addresses, get lost in busy inboxes, or simply be ignored or rejected [11,23]. Moreso, decisions to share are reliant on an individual’s subjective evaluation of what is “reasonable” unless clear criteria for requests are articulated and consistently applied.[24] If data can legally and safely be shared, it should be the default to allow for proactive transparency and ease of use.

The first article in this series covered pre-registration which can be considered a specific type of sharing of analysis plans and protocols prior to study commencement.[16] Making these available allows for assessments of potential biases and provides details for future researchers. Making the analytic code used to clean, process, and analyse study data in programs like R, STATA, or Python available can help catch errors, facilitate future research, and provide insights into how a study protocol was practically implemented.[4] Observational health researchers working with large datasets can share the codelists (e.g., ICD-10 codes) that are used to extract analysis populations and define variables of interest in order to avoid duplicating efforts and ensure consistency in analyses. Those involved in qualitative studies can make

study materials, like interview guides, available to provide insights into how data was collected. Repositories, such as the National Heart, Lung and Blood Institute Biologic Specimen and Data Repository, facilitate access to biospecimens and other clinical and epidemiologic data to benefit other researchers.

These are just some common examples of how sharing may occur and aid in promoting reproducibility, secondary analyses, and increased utility of publicly funded research outputs. There is even evidence that more open sharing practices leads to higher citation rates of a given work.[17] For any study, we believe it is valuable to reflect on what materials were key to its conduct and whether or not it can, and should, be shared. It may also be prudent to consider sharing to the extent possible prior to publication, rather than after publication, to ensure that reviewers, editors, and other interested parties can review study details for issues, errors, or biases before they enter the literature.[18]

How can materials be shared?

Proponents of better data sharing have developed the FAIR standards for “Findability, Accessibility, Interoperability, and Reuse.”[19] The details of FAIR are specific to making data available and usable, and have recently been expanded to research software[20], but provide a useful framework for thinking about what makes sharing successful more generally:

- *Findability and Accessibility:* Sharing in obscure or unindexed locations impacts whether others can discover and use the information. Additionally, one-off sharing on personal, or other non-persistent, repositories does not guarantee interested parties will be able to access the information into the future. Where relevant, ensuring appropriate metadata is attached to your information will also allow for more efficient discovery by others.
- *Interoperability and Reuse:* Posting information in a repository without any documentation or instructions can mean these materials are indecipherable to others.

While sharing a protocol and statistical analysis plan that explains the methods and

analysis in detail is a minimum standard, further clarifications, possibly including a data dictionary; comment and document code; or share instructions or guides about data extraction can help the inspection and reuse of data. Some journals offer researchers the opportunity to publish formal descriptions of available datasets to increase their visibility and document their contents and use (e.g., Nature Scientific Data). Properly presenting and documenting data or code will also make their future use more effective and efficient.[34–36]. It can be useful to share information in standard formats that can be easily utilised across a variety of platforms. For instance, make your data available as a .CSV file rather than a program specific data file (e.g., a STATA .DTA file). Some industries or research areas will have niche standards for interoperable, open data formats (e.g., Clinical Data Interchange Standards Consortium standards). When sharing in non-standard or proprietary formats is the best option, it should be clearly documented how and what is necessary to access and use the information.

Specific methods for sharing will also depend on what is being shared. The simplest form of sharing would involve making any study materials available as supplements to the manuscript. This has the clear benefit of being immediately available to readers, however this practice may have other limitations. Links to files can break in journal redesigns and there may be restrictions as to how information can be shared, accessed, and reused. For instance, sharing a large dataset as a PDF would compromise its usability by others.

Alternatively, information can be shared on services and repositories created to persistently house academic outputs. Preprint servers offer the opportunity to share additional study materials, with few limitations, alongside an open access version of the manuscript. The Open Science Framework, Figshare, and Zenodo are some examples of more general purpose repositories that offer generous free storage of nearly any type of file, structured metadata, version control, and digital identifiers (e.g., DOIs) for easy citation and linkage and often integrate with one another (Box 3). More specialised repositories, like GitHub for code, offer

similar functionality and can be valuable tools for sharing. Using these repositories makes sharing as easy as putting a persistent URL or DOI into a manuscript that points to all available materials.

Some platforms are specifically designed for securely warehousing more sensitive data, particularly relevant in health and medical research. Vivli and YODA are two examples of repositories that help securely store and manage access to individual patient data from clinical research.[25] There is also increased investment in Trusted Research Environments in which data providers offer approved researchers firewalled access to conduct analysis in which the raw data would never leave the environment.[26] By sharing code designed for analysis within a given Trusted Research Environment, any researcher with access should be able to replicate, verify, or expand on a finding. It may also be worthwhile to check the requirements and recommended resources or repositories from the researcher's institution or funders for depositing and managing access to sensitive information.

Box 3: Examples of Sharing Platforms

Platform	URL
Open Science Framework (OSF)	https://osf.io/
Figshare	https://figshare.com/
Zenodo	https://zenodo.org/
Dryad	https://datadryad.org/
Octopus	https://octopuspublishing.org/
National Heart, Lung and Blood Institute	https://biolincc.nhlbi.nih.gov/home/

(NHLBI) Biologic Specimen and Data Repository	
GitHub	https://github.com/
Yale University Open Data Access (YODA)	https://yoda.yale.edu/
Vivli	https://vivli.org/

Barriers and precautions

While increased sharing may be an ideal, numerous barriers exist, especially for health and medical research.[27] The biggest concern with sharing information, especially data, from health and medical studies is the potential to expose sensitive or confidential information. Legal regimes, such as the General Data Protection Regulation (GDPR) in Europe, must also be considered when sharing sensitive data derived from research. A key component of sharing is that it must be done thoughtfully and responsibly.[28] As discussed above, there are specialised repositories which can help manage access to sensitive data.

For some studies, it may be prudent to make data sharing part of the initial participant consent process. Using this model, study participants would understand exactly what data they are providing and the ways in which it could be used or shared. If the risks of re-identification are well-described, understood and fully consented to, this may allow for more effective sharing of data.[29] Even when sharing data in accepted or responsible ways, care must be taken to ensure information is not overly-disclosive or the anonymity of the data is over-emphasized. Pseudonymisation is a common practice in which some identifiable data is removed, however this does not mean that the data is not disclosive.[30] Furthermore, thought should be given to whether existing datasets could be combined with your open data to de-anonymise participants.[31] Data leakage and vulnerabilities may even exist in unexpected places. For

instance, SVG files, a common image format requested by journals, embeds the data underlying the image directly in the file and could be disclosive.

Researchers should consider ownership or permission to legally share data. If not possible, they should be transparent about the reasons for not sharing and who owns and manages the data and how they used this dataset. For instance, one might do a study using the UK Clinical Practice Research Datalink (CPRD) dataset which is built from primary care electronic health record data, or with proprietary data purchased from a commercial data supplier such as those detailing insurance claims or drug sales. Nevertheless, other materials could still be shared (e.g., the protocol and analytic code). When these considerations are not relevant, information can be shared under appropriately permissive licensing like Creative Commons or The MIT Licence for software.

Lastly, properly sharing data requires both human and financial resources to build data sharing capacity, access platforms, respond to inquiries and meeting data security specifications. Moreover, the right incentives might promote data sharing initiatives [33].

Conclusions

Increased transparency into the methods underlying published research is growing in importance. Concerns about rigour, reproducibility, and research integrity cannot be addressed if access to the key ingredients that underlie published analyses are unavailable. Institutions throughout biomedical research need to invest in educating researchers about sharing while establishing infrastructure, incentives, and requirements that support making study materials as accessible as possible.

Contributions

NJD wrote the first draft of the piece. All authors contributed to the conceptualisation of the manuscript and critically reviewed and provided feedback on its content.

Competing Interests

GCR was financially supported by the NHS, National Institute of Health Research (NIHR), School for Primary Care Research (SPCR), the Naji Foundation and the Rotary Foundation to study for a Doctor of Philosophy (DPhil) at the University of Oxford (2017-2020), but no longer has interests to declare. GCR is an Associate Editor of BMJ Evidence Based Medicine. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The remaining authors declare no competing interests relevant to this piece.

Data Availability Statement

Not Applicable.

Patient and Public Involvement

Patients and the broader public were not involved in the development of this piece.

References

- 1 Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996;**313**:570–1.
- 2 Lawrence JM, Meyerowitz-Katz G, Heathers JAJ, *et al.* The lesson of ivermectin: meta-analyses based on summary data alone are inherently unreliable. *Nat Med* 2021;**27**:1853–4.
- 3 Ross JS, Lehman R, Gross CP. The Importance of Clinical Trial Data Sharing. *Circ Cardiovasc Qual Outcomes* 2012;**5**:238–40.
- 4 Goldacre B, Morton CE, DeVito NJ. Why researchers should share their analytic code. *BMJ* 2019;**367**. doi:10.1136/bmj.l6365
- 5 Munafò MR, Nosek BA, Bishop DVM, *et al.* A manifesto for reproducible science. *Nature Human Behaviour* 2017;**1**:0021.
- 6 Jorgenson LA, Wolinetz CD, Collins FS. Incentivizing a New Culture of Data Stewardship: The NIH Policy for Data Management and Sharing. *JAMA* 2021;**326**:2259–60.

- 7 NIHR position on the sharing of research data. National Institute for Health and Care Research. 2021.<https://www.nihr.ac.uk/documents/nihr-position-on-the-sharing-of-research-data/12253> (accessed 20 Jul 2022).
- 8 Tedersoo L, Küngas R, Oras E, *et al.* Data sharing practices and data availability upon request differ across scientific disciplines. *Sci Data* 2021;**8**:192.
- 9 Chan A-W, Song F, Vickers A, *et al.* Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;**383**:257–66.
- 10 Taichman DB, Sahni P, Pinborg A, *et al.* Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors. *Ann Intern Med* 2017;**167**:63–5.
- 11 Gabelica M, Bojčić R, Puljak L. Many researchers were not compliant with their published data sharing statement: mixed-methods study. *J Clin Epidemiol* Published Online First: 30 May 2022. doi:10.1016/j.jclinepi.2022.05.019
- 12 Bergeat D, Lombard N, Gasmi A, *et al.* Data Sharing and Reanalyses Among Randomized Clinical Trials Published in Surgical Journals Before and After Adoption of a Data Availability and Reproducibility Policy. *JAMA Netw Open* 2022;**5**:e2215209.
- 13 Lo B. Sharing clinical trial data: maximizing benefits, minimizing risk. *JAMA* 2015;**313**:793–4.
- 14 Nosek BA, Alter G, Banks GC, *et al.* Promoting an open research culture. *Science* 2015;**348**:1422–5.
- 15 Eklund A, Nichols TE, Knutsson H. Reply to Brown and Behrmann, Cox, *et al.*, and Kessler *et al.*: Data and code sharing is the way forward for fMRI. *Proc. Natl. Acad. Sci. U. S. A.* 2017;**114**:E3374–5.
- 16 Cashin AG, Richards GC, DeVito NJ, *et al.* Registration of health and medical research. *BMJ Evid Based Med* Published Online First: 21 December 2021. doi:10.1136/bmjebm-2021-111836
- 17 Colavizza G, Hrynaszkiewicz I, Staden I, *et al.* The citation advantage of linking publications to research data. *PLoS One* 2020;**15**:e0230416.
- 18 Heyard R, Held L. When should data and code be made available? *Signif (Oxf)* 2022;**19**:4–5.
- 19 Stall S, Yarmey L, Cutcher-Gershenfeld J, *et al.* Make scientific data FAIR. Nature Publishing Group UK. 2019. doi:10.1038/d41586-019-01720-7
- 20 Chue Hong NP, Katz DS, Barker M, *et al.* FAIR Principles for Research Software (FAIR4RS Principles). 2022.<https://zenodo.org/record/6623556>
- 21 McGuinness LA, Sheppard AL. A descriptive analysis of the data availability statements accompanying medRxiv preprints and a comparison with their published counterparts. *PLoS One* 2021;**16**:e0250887.
- 22 DeBlanc J, Kay B, Lehrich J, *et al.* Availability of Statistical Code From Studies Using

Medicare Data in General Medical Journals. *JAMA Intern Med* Published Online First: 13 April 2020. doi:10.1001/jamainternmed.2020.0671

- 23 Siebert M, Gaba J, Renault A, *et al.* Data-sharing and re-analysis for main studies assessed by the European Medicines Agency-a cross-sectional study on European Public Assessment Reports. *BMC Med* 2022;**20**:177.
- 24 Sydes MR, Johnson AL, Meredith SK, *et al.* Sharing data from clinical trials: the rationale for a controlled access approach. *Trials* 2015;**16**:104.
- 25 Vazquez E, Gouraud H, Naudet F, *et al.* Characteristics of available studies and dissemination of research using major clinical data sharing platforms. *Clin Trials* 2021;;17407745211038524.
- 26 Goldacre B, Morley J, Hamilton N. Better, Broader, Safer: Using Health Data for Research and Analysis. Department of Health and Social Care 2022.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1067053/goldacre-review-using-health-data-for-research-and-analysis.pdf
- 27 Hopkins C, Sydes M, Murray G, *et al.* UK publicly funded Clinical Trials Units supported a controlled access approach to share individual participant data but highlighted concerns. *J Clin Epidemiol* 2016;**70**:17–25.
- 28 Lowenberg D, Puebla I. Responsible handling of ethics in data publication. *PLoS Biol* 2022;**20**:e3001606.
- 29 Garrison N 'a, Sathe NA, Antommaria AHM, *et al.* A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. *Genet Med* 2016;**18**:663–71.
- 30 Culnane C, Rubinstein BIP, Teague V. Health Data in an Open World. arXiv [cs.CY]. 2017.<http://arxiv.org/abs/1712.05627>
- 31 Simon GE, Shortreed SM, Coley RY, *et al.* Assessing and Minimizing Re-identification Risk in Research Data Derived from Health Care Records. *EGEMS (Wash DC)* 2019;**7**:6.
- 32 Popkin G. Data sharing and how it can benefit your scientific career. *Nature* 2019;**569**:445–7.
- 33 Rowhani-Farid A, Allen M, Barnett AG. What incentives increase data sharing in health and medical research? A systematic review. *Res Integr Peer Rev* 2017;**2**:4.
- 34 Barnes N. Publish your computer code: it is good enough. *Nature* 2010;**467**:753.
- 35 Valley TS, Kamdar N, Wiitala WL, *et al.* Continuous quality improvement in statistical code: avoiding errors and improving transparency. *BMJ Qual Saf* 2021;**30**:240–4.
- 36 Morton C, Devito N, Morley J, *et al.* Software development skills for health data researchers. *BMJ Health Care Inform* 2022;**29**. doi:10.1136/bmjhci-2021-100488
- 37 Nüst D, Sochat V, Marwick B, *et al.* Ten simple rules for writing Dockerfiles for reproducible data science. *PLoS Comput Biol* 2020;**16**:e1008316.

- 38 Wallach JD, Boyack KW, Ioannidis JPA. Reproducible research practices, transparency, and open access data in the biomedical literature, 2015-2017. *PLoS Biol* 2018;**16**:e2006930.
- 39 Serghiou S, Contopoulos-Ioannidis DG, Boyack KW, *et al.* Assessment of transparency indicators across the biomedical literature: How open is open? *PLoS Biol* 2021;**19**:e3001107.
- 40 Cashin AG, Bagg MK, Richards GC, *et al.* Limited engagement with transparent and open science standards in the policies of pain journals: a cross-sectional evaluation. *BMJ Evid Based Med* 2021;**26**:313–9.
- 41 Gardener AD, Hicks EJ, Jacklin C, *et al.* Open science policies of medical and health sciences journals before and during the COVID-19 pandemic: a repeat cross-sectional study. *bioRxiv*. 2022. doi:10.1101/2022.01.26.22269868
- 42 Ohmann C, Moher D, Siebert M, *et al.* Status, use and impact of sharing individual participant data from clinical trials: a scoping review. *BMJ Open* 2021;**11**:e049228.
- 43 Tan AC, Askie LM, Hunter KE, *et al.* Data sharing-trialists' plans at registration, attitudes, barriers and facilitators: A cohort study and cross-sectional survey. *Res Synth Methods* 2021;**12**:641–57.
- 44 DeVito NJ, French L, Goldacre B. Noncommercial Funders' Policies on Trial Registration, Access to Summary Results, and Individual Patient Data Availability. *JAMA* 2018;**319**:1721–3.
- 45 Whitlock EP, Dunham KM, DiGioia K, *et al.* Noncommercial US Funders' Policies on Trial Registration, Access to Summary Results, and Individual Patient Data Availability. *JAMA Netw Open* 2019;**2**:e187498.
- 46 Gaba JF, Siebert M, Dupuy A, *et al.* Funders' data-sharing policies in therapeutic research: A survey of commercial and non-commercial funders. *PLoS One* 2020;**15**:e0237464.
- 47 Axson S, Mello MM, Lincoff D, *et al.* Clinical trial transparency and data sharing among biopharmaceutical companies and the role of company size, location and product type: a cross-sectional descriptive analysis. *BMJ Open* 2021;**11**:e053248.
- 48 Goldacre B, Lane S, Mahtani KR, *et al.* Pharmaceutical companies' policies on access to trial data, results, and methods: audit study. *BMJ* 2017;**358**:j3334.
- 49 Siebert M, Gaba JF, Caquelin L, *et al.* Data-sharing recommendations in biomedical journals and randomised controlled trials: an audit of journals following the ICMJE recommendations. *BMJ Open* 2020;**10**:e038887.
- 50 Naudet F, Sakarovich C, Janiaud P, *et al.* Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in The BMJ and PLOS Medicine. *BMJ* 2018;**360**:k400.
- 51 Seibold H, Czerny S, Decke S, *et al.* A computational reproducibility study of PLOS ONE articles featuring longitudinal data analyses. *PLoS One* 2021;**16**:e0251194.
- 52 Page MJ, Altman DG, Shamseer L, *et al.* Reproducible research practices are underused in

- systematic reviews of biomedical interventions. *J Clin Epidemiol* 2018;**94**:8–18.
- 53 Stodden V, Seiler J, Ma Z. An empirical analysis of journal policy effectiveness for computational reproducibility. *Proc Natl Acad Sci U S A* 2018;**115**:2584–9.
 - 54 Besançon L, Peiffer-Smadja N, Segalas C, *et al.* Open science saves lives: lessons from the COVID-19 pandemic. *BMC Med Res Methodol* 2021;**21**:117.
 - 55 Lowe D. Surgisphere and Their Data. Science. 2020.<https://www.science.org/content/blog-post/surgisphere-and-their-data-updated> (accessed 5 May 2022).
 - 56 Weissgerber T, Riedel N, Kilicoglu H, *et al.* Automated screening of COVID-19 preprints: can we help authors to improve transparency and reproducibility? *Nat Med* 2021;**27**:6–7.
 - 57 Kadakia KT, Beckman AL, Ross JS, *et al.* Leveraging Open Science to Accelerate Research. *N Engl J Med* Published Online First: 24 March 2021. doi:10.1056/NEJMp2034518