medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Adherence to data-sharing policies - a comparison of the BMJ with other major medical journals

Alison Avenell[1] (Orcid: 0000-0003-4813-5628) & Dorothy V.M. Bishop[2] (Orcid: 0000-0002-2448-4033)

[1]Aberdeen Centre for Evaluation, University of Aberdeen, Foresterhill, Scotland
[2]Department of Experimental Psychology, University of Oxford, UK

Abstract

Background: In 2024, the BMJ updated its data-sharing policy for clinical trials, requiring deidentified individual patient data (IPD) to be openly deposited prior to publication. Our objective was to discover if data-sharing increased after introduction of the new policy.
Method: All data-sharing statements were downloaded from BMJ trials published in 2023 (submitted pre-updated policy) and 2025 (submitted post-updated policy). Data for 2025 were gathered for trials in five comparison medical journals. Data-sharing statements were coded to specify whether IPD were immediately available, and if not, the reason why. Where a statement gave a link to a repository, we checked whether data were available.
Results: Openly available IPD for BMJ trials increased from 0/32 prior to the new policy to 19/33 (58%) after the updated policy; seven articles gave repository links that did not yield any data. In the five comparison journals, rates of open IPD varied from 0% to 5.6%.
Conclusions: There was a substantial increase in open sharing of IPD after introduction of the new policy compared to a prior period. Open sharing of IPD is possible, but it is unpopular with authors and is unlikely to be achieved without firm editorial enforcement.

(195 words)

KEYWORDS: open data, clinical trials, medical journals

Introduction

In 2016, the International Committee of Medical Journal Editors (ICMJE) put forward a proposal for promoting data sharing for clinical trials, which was published simultaneously in Annals of Internal Medicine, British Medical Journal, Canadian Medical Association Journal, Chinese Medical Journal, Deutsches Ärzteblatt, Ethiopian Journal of Health Sciences, JAMA (Journal of the American Medical Association), Nederlands Tijdschrift voor Geneeskunde, New England Journal of Medicine (NEJM), New Zealand Medical Journal, PLOS Medicine, Revista Médica de Chile, The Lancet, and Ugeskrift for Laeger (Taichman et al, 2016). This stated that researchers had an ethical obligation to "to share with others the deidentified individual-patient data (IPD) underlying the results presented in the article (including tables, figures, and appendices or supplementary material) no later than 6 months after publication". In response to feedback, a modified recommendation was published in 2017 (Taichman et al., 2017), which required clinical trials to have a data-sharing statement, but fell short of mandating open IPD, while continuing to "envision a global research community in which sharing deidentified data becomes the norm." Two new journals (Bulletin of the World

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Health Organisation, and Korean Journal of Medical Science) were signatories to the 2017 recommendation - but three of the original signatories (Canadian Medical Association Journal, Chinese Medical Journal and Nederlands Tijdschrift voor Geneeskunde) were absent; it was unclear whether they found the new recommendation too onerous, too weak, or whether there were other reasons for not signing.

Given that there was much debate about the original ICMJE recommendations, it is rather surprising that the organisation said very little about *why* it is good to share IPD. Taichman et al (2017) stated: "*The International Committee of Medical Journal Editors (ICMJE) believes there is an ethical obligation to responsibly share data generated by interventional clinical trials because trial participants have put themselves at risk", before going to mention "substantial challenges*" that precluded mandating sharing of IPD "*at this time*".

Smith and Roberts (2016) outlined seven reasons that were given for not sharing data, declaring them all to be invalid. These were: (i) others using the data without having gone to the trouble of collecting them; (ii) being scooped; (iii) conclusions may not replicate; (iv) an alternative analysis could challenge published results; (v) sharing data may reveal weaknesses in data management; (vi) concerns about patient confidentiality; and (vii) technical reasons. Nevertheless, the optimistic title of Smith and Roberts' opinion piece, "Time for sharing data to become routine" has not been fulfilled.

Danchev et al (2021) looked at trials published in JAMA, Lancet and NEJM after they had adopted the ICMJE policy, and found that whereas 334/487 articles had a statement about data-sharing, only two IPD datasets were deidentified and openly available. Bergeat et al (2022) considered data-sharing for RCTs published in ten leading surgical journals, comparing 65 studies published before the policy was implemented and 65 published after. A data-sharing statement was provided for 11 of 65 RCTs after the policy versus none before the policy, but actual sharing was rare - IPD were available for two studies prior to the policy and two after the policy. Hamilton et al (2023) performed a systematic review of articles that evaluated data-sharing of medical studies in the period 2016-2021, and estimated that the proportion stating they would openly share IPD was 8%, whereas the proportion who actually shared was 2%.

The BMJ went beyond the ICMJE recommendations in introducing a requirement for sharing of IPD on request for clinical trials in 2015 (Loder & Groves, 2015). In 2024, the BMJ strengthened its data-sharing policy to make sharing of IPD and code mandatory prior to publication for RCTs (Loder et al., 2024).

The value of data-sharing was subsequently reiterated in an editorial by the Editor-In-Chief (Abbasi, 2025), after serious questions were raised about the integrity of data underpinning a trial of stem cell injection to prevent heart failure after myocardial infarction, leading to the trial's retraction (Attar et al., 2025). Scrutiny of the open peer review documents showed that IPD were deposited in an open repository only after two requests by the editor, and after peer review had been conducted. Abbasi stated "*One purpose of mandatory data sharing is that it might act as a deterrent. This would now seem to be a false hope. A reasonable conclusion from this episode is that transparency and data sharing are not enough to create trust. We're therefore examining how to prevent a similar situation arising again. It's clear that greater scrutiny of the underlying data is required before publication."*

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

This case piqued our interest in the extent to which fully open data from clinical trials can be achieved if a policy of mandatory data-sharing is taken seriously by editors, so that authors understand they cannot simply ignore it. Abbasi (2025) noted that BMJ had refused to publish papers where authors were unable or unwilling to share data. This provides an interesting natural experiment, allowing us to compare the extent of data-sharing for papers submitted before and after the introduction of the more stringent data-sharing policy in May 2024. For comparison, we assessed data-sharing at five other prominent journals that publish clinical RCTs (Annals of Internal Medicine, JAMA, Nature Medicine, New England Journal of Medicine, and The Lancet), four of whom had been signatories of the 2017 statement on data sharing, and all of which have had policies regarding data-sharing or data-sharing statements (see Table 1). As well as considering rates of open data-sharing, we did an ad hoc qualitative analysis of the content of data-sharing statements, to gain some insight into reasons given for data availability restrictions when data are not shared.

(Table 1 about here)

Aims

To examine:

1. How often clinical trial IPD are shared in journals that ask authors to provide data-sharing statements?
2. Whether mandating openly accessible IPD is associated with an increase in trials providing IPD?
3. Where IPD are not available, what conditions do authors specify to allow access?

Method

For the six journals of interest (BMJ, Annals of Internal Medicine, JAMA, Nature Medicine, NEJM, and The Lancet), RCTs in each journal were identified by a search of the Dimensions database for articles with the keywords "trial" or "RCT" for the year of 2025. For the BMJ the same search was conducted for the period 1st January 2023 to 30th April 2024, before the more stringent data-sharing policy was introduced (1st May 2024). We considered also searching the other journals for the earlier period, but they had had no change to data sharing policies.

The authors scrutinised the resulting list of articles to exclude review articles, meta-analyses, or other non-RCT articles, leading to the dataset summarised in Table 2.

Our primary interest was in whether IPD were openly available. We found that it was necessary to check this to the point of locating the dataset, because there were cases where links were provided to a repository but data could not be found, or were inaccessible. In a few cases, a subset of data was openly deposited; these articles were coded as having data available if deposited data were sufficient to repeat the analysis in principle. (An example would be a case where demographic data were not open, but data were provided for primary and secondary outcomes).

We did not attempt to replicate findings reported in the articles, or to evaluate the quality and integrity of deposited data.

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

For articles without open IPD, we coded data-sharing statements according to whether they indicated that sharing of data was possible (coded "maybe"), or explicitly stated that data would not be shared (coded "no"), and recorded whether a data-sharing platform was used. We also devised an ad hoc coding scheme to categorise statements in terms of whether there were time-based or other restrictions on data access, or if specific reasons were given to justify not sharing. One author coded all articles using this scheme. Its reliability was checked by two other researchers independently analysing 100 of the 584 articles.

Results

Reliability of the coding scheme for data sharing statements.

For the basic 3-way categorisation for data being available (yes), possibly available (maybe), or not available (no), agreement was near perfect for all pairs of three raters (97%, 98% and 99%). For the "maybe" category, we initially attempted to distinguish statements that specified a delay period prior to proposals being considered for approval, and those that did not, but the language used in the data-sharing statements was often unclear, leading to poor agreement. We therefore merged these categories to give the broad category of "Sharing if conditions met". The pairwise agreement for the more granular codes shown in Table 2 was still imperfect at 81%, 86% and 84%, but we felt this was adequate to give a more qualitative impression of the relative frequency of different conditions that researchers specified for data-sharing. For the final categorisation of data-sharing, we used consensus codes for articles that had been multiply-coded, and the coding by one author (DB) for the remainder.

Rates and conditions of data-sharing

Results are shown in Table 2.

(Table 2 about here)

Our first question was how often clinical trial IPD are shared in journals that ask authors to provide data-sharing statements. It is clear that if authors are asked to provide data-sharing statements, they do so, but these very seldom indicate that IPD will be shared without any restrictions.

There was a pronounced increase in the proportion of articles with open IPD in the BMJ since introduction of mandatory data-sharing in 2024 (Fisher exact test comparing the pre- vs post-BMJ data, 0/32 (0%) v 19/33 (57.6%), $p < .001$). Rates of open IPD in BMJ prior to the policy are comparable to those seen in the comparison journals in 2025, which ranged from 0% to 5.6%.

There were seven BMJ trials published after the BMJ changed its policy where we could not access data, despite statements to the contrary. Four trials had links to the Dryad repository. Correspondence with Dryad revealed that data from these trials had never been deposited. The BMJ was notified of missing data from these trials on 28th February 2026. As of 11th May 2026 no revisions or notices had been posted by the BMJ for these trials, and data remained unavailable.

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

For the comparison journals, where a data-sharing statement was provided, it was relatively uncommon for authors to state they would share data "*on (reasonable) request*". A typical data-sharing statement would request that researchers submit a proposal for approval and require a data-sharing agreement to be signed if approval was given. Frequently, there were additional conditions to be met, concerning the time frame of requests (often only possible after a delay), the qualifications of the research team, and/or approvals from ethics committees. Some data-sharing statements were too vague to be useful.

Use of data-sharing platforms was relatively uncommon, but was seen particularly for studies sponsored by pharmaceutical companies, where the Vivli platform was the most popular (see Supplementary Table 1). Vivli is one of several data-sharing platforms that were set up to encourage data reuse, while controlling access to ensure responsible use. Requests for data are submitted via a standard form, and evaluated by an independent committee. Users are required to sign a Data Use Agreement.

Discussion

Study limitations

This study relies on natural observational data, so cannot prove the policy change caused the increase in data-sharing seen in the BMJ. Furthermore, the coding used in Table 2 was devised by a process of trial and error to settle on categories that were both reasonably reliable and provided useful information, but its reliability is not established in a new dataset. Using a standardised form to collect details, for example, as undertaken by the NEJM, would likely remove some of the ambiguities in description which challenged us in our coding, providing that all questions were answered.

We did not look for open code, nor did we examine whether reported results could be replicated from IPD.

Data sharing higher when mandated

These results suggest that, despite frequent statements by journal editors, drug companies, and researchers about the value of open data, sharing of IPD is strongly resisted by most researchers. The case of the BMJ is important in suggesting that open data can be achieved if it is a condition for publishing in the journal, but this policy needs to be enforced by editors. Furthermore, some authors gave the impression of compliance by including links to a data repository but then failed to deposit their data. Journal editors need to check that open data exists, rather than relying on statements to that effect. At minimum, random journal checks are needed to confirm that data are available when the authors give a link to a repository.

Reasons for widespread resistance to data sharing

The question arises as to why there is such strong resistance to data-sharing. In a minority of cases (ranging from 0% to 11.8%) data-sharing statements simply said data would not be shared, without giving any reason. A similar proportion justified not sharing because of ethical, legal or commercial considerations. With the exception of BMJ 2025, the most common type of data sharing statement was one where it was noted that sharing was possible if certain conditions were met. Typically, these required those using data to submit a proposal for approval and to sign a data-sharing agreement.

The impression was that authors had concerns similar to those outlined by Smith and Roberts (2016), especially inappropriate use of data, allowing others to take the credit for the work, or being scooped. These extracts from data-sharing statements illustrate these points:

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

L033: "*Gilead Sciences shares anonymised individual participant data upon request or as required by law or regulation with qualified external researchers based on submitted curriculum vitae and reflecting non-conflict of interest. The request proposal must also include a statistician. Approval of such requests is at Gilead Science's discretion and is dependent on the nature of the request, the merit of the research proposed, the availability of the data, and the intended use of the data."*
N033: (In response to question When will data availability begin?) "*After the medicine and indication is approved by both the FDA and EMA*."
N061: "*Investigators who have invested time and effort into developing a trial or study should have a period of exclusivity in which to pursue their aims with the data, before key trial data are made available to other researchers."*
A018: "*The PHRI believes the dissemination of clinical research results is vital and sharing of data is important. PHRI prioritizes access to data analyses to researchers who have worked on the trial for a substantial duration, have played substantial roles, and have participated in raising the funds to conduct the trial. PHRI balances the length of the research study, and the intellectual and financial investments that made it possible with the need to allow wider access to the data collected. Data will be disclosed only upon request and approval of the proposed use of the data by a Review Committee. Data will be available to the journal for evaluation of reported analyses. Data requests from other non-POISE-3 investigators will not be considered until 5 years after the closeout of the trial*."
Another concern that was sometimes explicitly stated was that providing data would be time-consuming for the original research group:
M090:"*Data requests will be considered by the Sponsor and the University of Glasgow, which will take account of the scientific rationale, ethics, logistics and resource implications....Requests for transfer of deidentified data (including source imaging scans) will be considered by the University of Glasgow, and, if approved, a collaboration agreement would be expected, taking account of any cost implications. Cost recovery would be expected on a not-for-profit basis."*

Trials sponsored by pharmaceutical companies often gave a generic data-sharing statement. This reflected a commitment to data-sharing, in the context of needing to protect commercial interests and adhere to ethical standards, as well as preventing use of the data for unmotivated data-dredging. This statement from a study sponsored by Merck & Co illustrates these points:
M113: "*MSD is committed to providing qualified scientific researchers access to anonymized data and clinical study reports from the company's clinical trials for the purpose of conducting legitimate scientific research.... Feasible requests will be reviewed by a committee of MSD subject matter experts to assess the scientific validity of the request and the qualifications of the requesters. In line with data privacy legislation, submitters of approved requests must enter into a standard data-sharing agreement with MSD before data access is granted. Data will be made available for request after product approval in the United States and the European Union or after product development is discontinued.....*"

Smith and Roberts (2016) were dismissive of such concerns, particularly where they appeared to reflect a "data-hoarding" approach, or involved barriers being erected to prevent the researchers from close scrutiny of their data management or datasets. In general, patients endorse sharing of their data, providing it is anonymised (National Academies of Sciences, Engineering and Medicine, 2020). Nevertheless, we know from numerous studies that, unless IPD are deposited with a permanent DOI, they are unlikely to ever be made available to other researchers, regardless of what the data-sharing statement says (Hussey, 2025;

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Kotoulas et al., 2026). Quite simply, websites decay and research groups disband and become uncontactable (Vines et al., 2014). Data-sharing is less onerous when planned for at the start of a project: reconstructing a dataset after the study is completed can involve much more work.

It is important to recognise that there are valid reasons for concerns about data misuse, as illustrated by recent examples of AI generation of pointless papers from large open datasets (Maupin et al., 2025), or use of cherrypicked data for ideological rather than scientific ends (Bishop, 2016). We are not aware of such cases affecting data from RCTs, but with the growth of easily-accessible methods for scraping open data, the potential danger exists. The challenge for journals and researchers is how to protect data from misuse without losing the benefits of open data-sharing.

Use of data-sharing platforms

As shown in Supplementary Table 1, some studies mentioned that data could be requested via independent data-sharing platforms. Such platforms were discussed at length in a workshop on clinical trial data sharing run by the US National Academies of Sciences, Engineering and Medicine (2020). Platforms such as Vivli use independent review panels to scrutinise requests for data-sharing, and require data-use agreements to ensure that those using data do so ethically. The benefits of such platforms are that they provide a transparent access route, overseen by an independent data committee, while preventing inappropriate use. However, they have disadvantages: first, their costs may not be justified by their apparent benefits, and it is unclear who should bear those costs; second, they can impose further delays and bureaucratic burdens on those seeking access to data; it was noted in the NASEM report that institutions often are slow to approve data-sharing agreements, and may wrangle over precise wording. For those who wish simply to check the accuracy of a deposited dataset, these barriers may seem like overkill.

The growing need to check trial data integrity

It is notable that very few of the conditional data-sharing statements or journal statements in our study showed any appreciation of the value and need for data-sharing to check the integrity of data and analyses. There is growing concern that the number of fraudulent clinical trials may be increasing, and it has been argued that sharing of IPD can help detect these (Van Noorden, 2023; Brown & Carlisle, 2025). The editor of the journal Anaesthesia (Carlisle, 2021), showed that checking IPD trial data flagged 26% of trial submissions as "zombie" trials, as opposed to 1% detection with trial level data. In a further study requesting IPD at submission to the journal (Bramley 2025), 87% of trial authors agreed to provide IPD, and 9% of authors reported finding errors in their data. Provision of IPD to journals for data integrity checks might be more acceptable to authors, but this would not constitute open access. Currently, many pharmaceutical researchers refuse to share IPD until product approval has been obtained; this makes sense from a commercial point of view, but it also implies that the data are error-free; in practice, if there are errors in the data, it would be preferable for patient safety to find them prior to product approval.

Prominent journals are not immune to fraudulent data. Two of the journals in our analysis (the Lancet and NEJM) published analyses based on fraudulent registry data from Surgisphere, and it was noted that identification of the issues was hindered by inability to check original data (Editors of the Lancet, 2020; Mehra et al., 2020). Our current analysis was motivated by findings of data irregularities in a trial published in the BMJ, which has since been retracted. It is unlikely that this trial would have been questioned if data had not

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

been openly available. Quite apart from such misconduct, there is ample evidence that honest errors in data and analysis are common, and unless we can detect them, they will remain to corrupt published research. Openly available data, without the need for permissions, would allow much faster identification and correction or retraction of studies with data irregularities.

It is 10 years since the original ICMJE statement. Although there has been some progress since, the landscape of trial integrity issues has been changed by increasing numbers of fraudulent studies (van Noorden, 2023). Journals may be forced to rethink their policies in order to retain public confidence in the trials that they publish.

Acknowledgements

We are most grateful to Mark Bolland for his helpful discussions about this article, and for contributing to data coding.

Funding

None.

Conflicts of interest

None to declare.

Ethical approval

Not applicable.

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

References

Abbasi, K. (2025). The potential and limits of scrutiny in medical research. *BMJ*, *391*, r2411. https://doi.org/10.1136/bmj.r2411

Attar, A., Mirhosseini, S. A., Mathur, A., Dowlut, S., Monabati, A., Kasaei, M., Abtahi, F., Kiwan, Y., Vosough, M., & Azarpira, N. (2025). Prevention of acute myocardial infarction induced heart failure by intracoronary infusion of mesenchymal stem cells: Phase 3 randomised clinical trial (PREVENT-TAHA8). *BMJ*, *391*, e083382. https://doi.org/10.1136/bmj-2024-083382

Bergeat, D., Lombard, N., Gasmi, A., Le Floch, B., & Naudet, F. (2022). Data sharing and reanalyses among randomized clinical trials published in surgical journals before and after adoption of a data availability and reproducibility policy. *JAMA Network Open*, *5*(6), e2215209. https://doi.org/10.1001/jamanetworkopen.2022.15209

Bishop, D. V. M. (2016). Open research practices: Unintended consequences and suggestions for averting them. (Commentary on the Peer Reviewers' Openness Initiative). *Royal Society Open Science*, *3*(4), 160109. https://doi.org/10.1098/rsos.160109

Bramley, P. (2025). Ask, and it shall be given you – individual patient data and code availability for randomised controlled trials submitted for publication. *Anaesthesia*, *80*(2), 205–206. https://doi.org/10.1111/anae.16503

Brown, N. J. L., & Carlisle, J. B. (2025). Detecting inconsistencies and fraud in research data: Time for authors to share the data underlying their summary statistics as a matter of course. *Anesthesia & Analgesia*, 10.1213/ANE.0000000000007889. https://doi.org/10.1213/ANE.0000000000007889

Carlisle, J. B. (2021). False individual patient data and zombie randomised controlled trials submitted to Anaesthesia. *Anaesthesia*, *76*(4), 472–479. https://doi.org/10.1111/anae.15263

Danchev, V., Min, Y., Borghi, J., Baiocchi, M., & Ioannidis, J. P. A. (2021). Evaluation of data sharing after implementation of the International Committee of Medical Journal Editors data sharing statement requirement. *JAMA Network Open*, *4*(1), e2033972. https://doi.org/10.1001/jamanetworkopen.2020.33972

Editors of the Lancet. (2020). Learning from a retraction. *The Lancet*, *396*(10257), 1056. https://doi.org/10.1016/S0140-6736(20)31958-9

Hamilton, D. G., Hong, K., Fraser, H., Rowhani-Farid, A., Fidler, F., & Page, M. J. (2023). Prevalence and predictors of data and code sharing in the medical and health sciences: Systematic review with meta-analysis of individual participant data. *BMJ*, *382*, e075767. https://doi.org/10.1136/bmj-2023-075767

Hussey, I. (2025). Data is not available upon request. *Meta-Psychology*, *9*. https://doi.org/10.15626/MP.2023.4008

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Kotoulas, B., Blackwell, J., & Hardwicke, T. E. (2026). *Has transparency improved in clinical psychology? A repeated cross-sectional study (2012, 2018, 2024)* (Jn435_v1). PsyArXiv. https://osf.io/preprints/psyarxiv/jn435_v1/

Loder, E., & Groves, T. (2015). The BMJ requires data sharing on request for all trials. *BMJ*, *350*, h2373. https://doi.org/10.1136/bmj.h2373

Loder, E., Macdonald, H., Bloom, T., & Abbasi, K. (2024). Mandatory data and code sharing for research published by the BMJ. *BMJ*, *384*, q324. https://doi.org/10.1136/bmj.q324

Maupin, D., Spick, M., & Geifman, N. (2025). Safeguarding Open Science from exploitative practices. *PLOS Medicine*, *22*(12), e1004851. https://doi.org/10.1371/journal.pmed.1004851

Mehra, M. R., Desai, S. S., Kuy, S., Henry, T. D., & Patel, A. N. (2020). Retraction: Cardiovascular disease, drug therapy, and mortality in Covid-19. N Engl J Med. DOI: 10.1056/NEJMoa2007621. | NEJM (Pt 2582). *New England Journal of Medicine*, *382*. https://www.nejm.org/doi/full/10.1056/NEJMc2021225

National Academies of Sciences, Engineering, and Medicine (2020). *Reflections on Sharing Clinical Trial Data: Challenges and a Way Forward: Proceedings of a Workshop* (C. Shore, J. Hinners, E. Khandekar, & T. Wizemann, Eds). National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK559433/

Smith, R., & Roberts, I. (2016). Time for sharing data to become routine: The seven excuses for not doing so are all invalid [version 1; peer review: 2 approved, 1 approved with reservations]. *F1000Research*, *5*, 781. https://doi.org/10.12688/f1000research.8422.1

Taichman, D. B., Backus, J., Baethge, C., Bauchner, H., Leeuw, P. W. de, Drazen, J. M., Fletcher, J., Frizelle, F. A., Groves, T., Haileamlak, A., James, A., Laine, C., Peiperl, L., Pinborg, A., Sahni, P., & Wu, S. (2016). Sharing clinical trial data: A proposal from the International Committee of Medical Journal Editors. *PLOS Medicine*, *13*(1), e1001950. https://doi.org/10.1371/journal.pmed.1001950

Taichman, D. B., Sahni, P., Pinborg, A., Peiperl, L., Laine, C., James, A., Hong, S.-T., Haileamlak, A., Gollogly, L., Godlee, F., Frizelle, F. A., Florenzano, F., Drazen, J. M., Bauchner, H., Baethge, C., & Backus, J. (2017). Data sharing statements for clinical trials: A requirement of the International Committee of Medical Journal Editors. *JAMA*, *317*(24), 2491–2492. https://doi.org/10.1001/jama.2017.6514

Van Noorden, R. (2023). Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed? *Nature*, *619*(7970), 454–458. https://doi.org/10.1038/d41586-023-02299-w

Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., Gilbert, K. J., Moore, J.-S., Renaut, S., & Rennison, D. J. (2014). The availability of research data declines rapidly with article age. *Current Biology*, *24*(1), 94–97. https://doi.org/10.1016/j.cub.2013.11.014

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Data sharing statement

The data and code used to create Table 2 are deposited at Figshare, DOI: 10.6084/m9.figshare.32270577.

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Table 1

Data-sharing policies of six leading medical journals

Journal	Data-sharing policy
BMJ - submitted pre 1 May 2024[1]	Share relevant trial data on reasonable request
BMJ - submitted post 1 May 2024[2]	For reports of clinical trials submitted to The BMJ on or after 1 May 2024, we require that the data underlying the results be available at the time of publication. This encompasses all anonymised data on individual patients on which the analyses, results, and conclusions reported in the paper are based. Data repositories should: • Have open licences or allow access without unreasonable barriers. • Provide a stable and persistent identifiers for datasets, such as a DOI.
Annals of Internal Medicine[3]	Clinical trials that begin enrolling participants on or after 1 January 2019 must include a data sharing plan in the trial's registration. If the data sharing plan changes after registration, this should be reflected in the statement submitted and published with the manuscript and updated in the registry record. Data sharing statements must indicate the following: whether individual deidentified participant data (including data dictionaries) will be shared ("undecided" is not an acceptable answer); what data in particular will be shared; whether additional, related documents will be available (e.g., study protocol, statistical analysis plan); when the data will become available and for how long; by what access criteria data will be shared (including with whom, for what types of analyses, and by what mechanism).
JAMA[4]	For all reports of research, authors are required to provide a Data Sharing Statement to indicate if data will or will not be shared. Specific questions regarding the sharing of data are included in the manuscript submission system. If authors choose to share or not share data, this information will be published in a Data Sharing Statement in an online supplement linked to the published article. Authors will be asked to identify the data, including individual patient data, a data dictionary that defines each field in the data set, and supporting documentation (eg, statistical/analytic code), that will be made available to others; when, where, and how the data will be available (eg, a link to a data repository); types of analyses that are permitted; and if there will be any restrictions on the use of the data. Authors also have the option to explain why data may not be shared.
The Lancet[5]	The Lancet Group is committed to open science and requires that all research manuscripts include a data sharing statement from the authors that details what

[1] BMJ: 2015 - 1 May 2024: Loder, E., & Groves, T. (2015). The BMJ requires data sharing on request for all trials. *BMJ*, *350*, h2373. https://doi.org/10.1136/bmj.h2373

[2] BMJ: post 1 May 2024: https://www.bmj.com/about-bmj/resources-authors/article-types/data-sharing

[3] Annals of Internal Medicine: https://www.acpjournals.org/pb-assets/pdf/AnnalsAuthorInfo-1775076005267.pdf

[4] JAMA: https://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecDataSharingStatement

[5] Lancet: https://www.thelancet.com/editorial-policies

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

	data will be shared, whether additional documents will be shared (eg, the study protocol), when data will become available, and by what access criteria data will be shared. A data sharing statement will be published in the article.
Nature Medicine[6]	All published manuscripts reporting original research in Nature Portfolio journals must include a data availability statement. The data availability statement must make the conditions of access to the "minimum dataset" that are necessary to interpret, verify and extend the research in the article, transparent to readers. This minimum dataset may be provided through deposition in public community/discipline-specific repositories, custom proprietary repositories for certain types of datasets, or general repositories like Figshare, Zenodo and Dryad. Providing large datasets in supplementary information is strongly discouraged and the preferred approach is to make data available in repositories
New England Journal of Medicine[7]	The International Committee of Medical Journal Editors (ICMJE) and, therefore, NEJM require investigators to submit a data-sharing statement (2018) and register a data-sharing plan when registering a trial (2019). NEJM has not yet formulated a policy on which types of data-sharing plans are acceptable.

[6] Nature Medicine: https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards

[7] NEJM: https://www.nejm.org/about-nejm/editorial-policies

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Table 2

Numbers (%) of articles categorized according to data-sharing statements

	BMJ-pre	BMJ-post	Annals of Internal Medicine	JAMA	Lancet	Nature Medicine	NEJM
N RCTs	32	33	18	76	111	154	160
Category of data-sharing statement							
Links to repository with IPD							
Link and data confirmed available	0 (0%)	19 (57.6%)	1 (5.6%)	1 (1.3%)	1 (0.9%)	8 (5.2%)	0 (0%)
Link but no data	0 (0%)	7 (21.2%)	0 (0%)	0 (0%)	1 (0.9%)	4 (2.6%)	7 (4.4%)
Will share in principle							
Sharing on request	9 (28.1%)	0 (0%)	3 (16.7%)	2 (2.6%)	3 (2.7%)	7 (4.5%)	12 (7.5%)
Sharing if conditions met (include requirement to submit proposal, sign datasharing agreement, get approvals, wait for specified duration)	19 (59.4%)	6 (18.2%)	10 (55.6%)	54 (71.1%)	74 (66.7%)	96 (62.3%)	71 (44.4%)
Will share after delay, no other conditions specified	1 (3.1%)	0 (0%)	0 (0%)	0 (0%)	2 (1.8%)	1 (0.6%)	4 (2.5%)
Will share but require collaboration	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (1.8%)	3 (1.9%)	1 (0.6%)
Will share but timing/conditions vague or uncertain	0 (0%)	0 (0%)	0 (0%)	4 (5.3%)	15 (13.5%)	23 (14.9%)	38 (23.8%)
Will not share							
No sharing - commercial/intellectual property	0 (0%)	1 (3%)	1 (5.6%)	1 (1.3%)	3 (2.7%)	2 (1.3%)	0 (0%)
No sharing - ethics/legal/privacy	1 (3.1%)	0 (0%)	2 (11.1%)	5 (6.6%)	7 (6.3%)	10 (6.5%)	7 (4.4%)
No sharing - other	1 (3.1%)	0 (0%)	0 (0%)	0 (0%)	2 (1.8%)	0 (0%)	4 (2.5%)
No sharing - no reason	1 (3.1%)	0 (0%)	1 (5.6%)	9 (11.8%)	1 (0.9%)	0 (0%)	15 (9.4%)
No statement	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.6%)

medRxiv preprint doi: https://doi.org/10.64898/2026.05.15.26353284; this version posted May 21, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Supplementary table 1

Data-sharing platforms mentioned by articles summarised in Table 2

	BMJ_pre 2025	BMJ 2025	Annals of Internal Medicine	JAMA	Lancet	Nature Medicine	NEJM
N RCTs	32	33	18	76	111	154	160
Platform							
None	28	7	15	66	86	118	101
Clinicaltrials.gov	1	0	0	0	0	1	0
CSDR	0	0	0	0	1	2	2
Dataverse	0	1	1	1	0	0	3
Dryad	0	6	0	0	0	0	0
Figshare	0	2	0	0	0	2	0
Mendeley	0	3	0	0	0	0	0
NCTN	0	0	0	0	0	0	3
Novonordisk	0	0	0	0	2	0	7
OSF	0	2	0	0	1	0	0
other	2	10	0	6	2	4	7
Research Data Deposit	1	1	0	0	0	1	1
Vivli	0	0	1	3	16	22	32
YODA	0	0	1	0	3	3	3
Zenodo	0	1	0	0	0	1	1
Other*	2	10	0	6	2	4	7

* other named platforms, each of which had 2 or less mentions