



Making data sharing the norm in medical research

The benefits to patients, science, and society are undeniable

Clara Locher,¹ Gérard Le Goff,² Anne Le Louarn,³ Ulrich Mansmann,⁴ Florian Naudet^{1,5}

¹ University of Rennes, CHU Rennes, Inserm, Irset (Institut de recherche en santé, environnement et travail)-UMR_S 1085, CIC 1414 (Centre of Clinical Investigation of Rennes, Rennes, France

² Patient representative, France Rein Bretagne, Lailé, France

³ GCS CNCR (Comité National de Coordination de la Recherche), Paris, France

⁴ Ludwig-Maximilians University Munich, Medical Faculty, Institute for Medical Information Processing, Biometry, and Epidemiology, Munich, Germany

⁵ University Institute of France (IUF), Paris, France

Correspondence to: F Naudet
florianaudet@gmail.com

Cite this as: *BMJ* 2023;382:p1434
<http://dx.doi.org/10.1136/bmj.p1434>

Published: 11 July 2023

Reuse of medical research data—which is conditional on access to individual participant data—is expected to maximise the value of medical research. It enables alternative hypothesis testing, validation of claims, exploration of controversies, restoration of unpublished trials, avoidance of duplicated efforts, and production of new knowledge from existing datasets. Given these benefits, politicians,^{1–3} funders,⁴ and publishers⁵ now support and implement data sharing policies. However, converging evidence indicates that current policies are unlikely to reach their goal of achieving data sharing. In a linked paper at *The BMJ*, Hamilton and colleagues (doi:10.1136/bmj-2023-075767) synthesised 105 meta-research studies examining 2 121 580 articles across 31 medical specialties and found that, despite some heterogeneity, data sharing rates are consistently low across medical research.⁶ Intention-to-share data have increased with time but are not associated with any increase in actual data sharing.

Responsible sharing of such sensitive data is, however, a complex endeavour. Preparing, storing, processing, administering data, and complying with legal and regulatory data protection requirements all call for appropriate infrastructure and institutional support, generating additional costs.⁷ While these practical problems can be resolved pre-emptively, the suboptimal sharing of systematic review data as shown in the study by Hamilton and colleagues⁶—which are neither sensitive nor difficult to share—suggests deeper reasons for the lack of open data.

Unlike the open source software community, where error correction is welcomed, sharing biomedical data can be perceived as risky given the possibility of uncovering errors that can damage a researcher's reputation. Some private companies and researchers can consider that they own the data they have collected and view data sharing as a competitive disadvantage via loss of the exclusive use of the data. Further, the value of reuse, replication, or validation studies is sometimes underestimated by editors.⁸

These barriers to data sharing focus more on individual researchers' or organisations' interests and less on expected benefits to science and society. We cannot overestimate the positive impact on citizens of knowing how their health data are reused, and the contribution to improving healthcare. Trial participants are generally supportive of responsible sharing of de-identified individual patient data.⁹ If data sharing is not without risk, the risk of re-identification of correctly pseudonymised information shared in a secure environment remains low in relation to the stakes. The new knowledge

produced through data sharing is likely to make patients more aware of the positive benefit/risk balance—including the benefit of improving health for all—thus refocusing the debate.

Coordinated efforts can improve the situation. If leading funders such as the US National Institutes of Health and Horizon Europe implement and finance data sharing activities, other funders should follow. The French DGOS (Direction Générale de l'Offre de Soins), for example, funds clinical research for about €130m (£112m; \$143m) per year¹⁰ but has no policy yet. Institutional review boards should systematically assess data sharing plans within their ethical evaluation of research projects, balancing the potential benefits (eg, large confirmatory trials might have more value than small exploratory studies) with the risks (eg, re-identification risks could be higher in trials of rare v frequent diseases). Institutional review boards should be able to assess in any new research project whether the research question would benefit from data sharing or whether it would be better resolved by reusing existing datasets, because data reuse avoids exposing new research participants to risk.

Journals and regulators should adopt mandatory policies, which are probably more effective than non-mandatory policies.⁶ The Committee on Publication Ethics should clarify how journals should react when confronted with empty promises of data-sharing. Initiatives rewarding data sharing behaviours, such as the Hong Kong Principles¹¹ or the Good Pharma Scorecard,¹² should be widely endorsed to establish openness and transparency as core values. But even if it is possible to make open data a reality, shared data are useless if they are not requested. Available datasets are clearly underused.¹³ There is a need to train researchers not only to prepare and share but also to reuse existing datasets.¹⁴ Such so-called data champions should disseminate the culture of open data everywhere (eg, funders, institutional review boards, data use and access committees), making data sharing a practice with real impact and not just wishful thinking.

Competing interests: The BMJ has judged that there are no disqualifying financial ties to commercial companies. The authors declare the following other interests: CL is involved in the doctoral network MSCA-DN SHARE-CTD (HORIZON-MSCA-2022-DN-01 101120360), funded by the EU, with additional contributions from Bayer, Smart Data Analysis and Statistics, SciCrunch, and ECRIN-ERIC. UM is the coordinator of MSCA-DN SHARE-CTD. FN received funding from the French National Research Agency (ANR-17-CE36-0010), the French Ministries of Health and Research. FN is a work package leader in the OSIRIS project (Open Science to Increase Reproducibility in Science); the OSIRIS project has received funding from the EU's Horizon Europe research an innovation programme. FN is also a WP leader in MSCA-DN SHARE-CTD. ALL and GLG no conflicts of interests. Further details of The BMJ policy on financial interests are here: <https://www.bmj.com/sites/default/files/attachments/resources/2016/03/16-current-bmj-education-coi-form.pdf>.

Provenance and peer review: Commissioned; not externally peer reviewed.

We thank Angela Swaine Verdier for reviewing the English language used in this article.

- 1 UNESCO. Recommendation on Open Science. <https://www.unesco.org/en/open-science/about>.
- 2 Tsukuba Communiqué. (G7 Science and Technology Ministers' Meeting in Tsukuba, Ibaraki). Cabinet Office Home Page. https://www8.cao.go.jp/cstp/english/others/communique_en.html.
- 3 European Federation of Pharmaceutical Industries and Associations. Principles for responsible clinical trial data sharing. 2013. <https://www.efpia.eu/media/25189/principles-for-responsible-clinical-trial-data-sharing.pdf>.
- 4 Gaba JF, Siebert M, Dupuy A, Moher D, Naudet F. Funders' data-sharing policies in therapeutic research: A survey of commercial and non-commercial funders. *PLoS One* 2020;15:e0237464. doi: 10.1371/journal.pone.0237464. pmid: 32817724
- 5 Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials. *BMJ* 2017;357. doi: 10.1136/bmj.j2372. pmid: 28584025
- 6 Hamilton D, Hong K, Fraser H, Rowhani-Farid A, Fidler F, Page MJ. Rates and predictors of data and code sharing in the medical and health sciences: systematic review with meta-analysis of individual participant data. *BMJ* 2023;382:e075767.
- 7 Ventresca M, Schünemann HJ, Macbeth F, et al. Obtaining and managing data sets for individual participant data meta-analysis: scoping review and practical guide. *BMC Med Res Methodol* 2020;20. doi: 10.1186/s12874-020-00964-6. pmid: 32398016
- 8 Longo DL, Drazen JM. Data Sharing. *N Engl J Med* 2016;374-7. doi: 10.1056/NEJMe1516564. pmid: 26789876
- 9 Mello MM, Lieou V, Goodman SN. Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing. *N Engl J Med* 2018;378-11. doi: 10.1056/NEJMs1713258. pmid: 29874542
- 10 Rollando P, Parc C, Naudet F, Gaba JF. [Data sharing policies of clinical trials funders in France] [French]. *Therapie* 2020;75:-36. doi: 10.1016/j.therap.2020.04.001. pmid: 32446662
- 11 Moher D, Bouter L, Kleinert S, et al. The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS Biol* 2020;18:e3000737. doi: 10.1371/journal.pbio.3000737. pmid: 32673304
- 12 Miller J, Ross JS, Wilenzick M, Mello MM. Sharing of clinical trial data and results reporting practices among large pharmaceutical companies: cross sectional descriptive study and pilot of a tool to improve company practices. *BMJ* 2019;366. doi: 10.1136/bmj.l4217. pmid: 31292127
- 13 Ohmann C, Moher D, Siebert M, Motschall E, Naudet F. Status, use and impact of sharing individual participant data from clinical trials: a scoping review. *BMJ Open* 2021;11:e049228. doi: 10.1136/bmjopen-2021-049228. pmid: 34408052
- 14 Mansmann U, Locher C, Prasser F, et al. Implementing clinical trial data sharing requires training a new generation of biomedical researchers. *Nat Med* 2023;29:-301. doi: 10.1038/s41591-022-02080-y. pmid: 36732626