CrossMark

RADIOLOGICAL EDUCATION

To share or not to share? Expected pros and cons of *data sharing* in radiological research

Francesco Sardanelli[1,2] · Marco Alì[3] · Myriam G. Hunink[4,5] · Nehmat Houssami[6] ·
Luca M. Sconfienza[1,7] · Giovanni Di Leo[2]

Abstract

The aims of this paper are to illustrate the trend towards *data sharing*, i.e. the regulated availability of the original patient-level data obtained during a study, and to discuss the expected advantages (pros) and disadvantages (cons) of data sharing in radiological research. Expected pros include the potential for verification of original results with alternative or supplementary analyses (including estimation of reproducibility), advancement of knowledge by providing new results by testing new hypotheses (not explored by the original authors) on pre-existing databases, larger scale analyses based on individual-patient data, enhanced multidisciplinary cooperation, reduced publication of false studies, improved clinical practice, and reduced cost and time for clinical research. Expected cons are outlined as the risk that the original authors could not exploit the entire potential of the data they obtained, possible failures in patients' privacy protection, technical barriers such as the lack of standard formats, and possible data misinterpretation. Finally, open issues regarding data ownership, the role of individual patients, advocacy groups and funding institutions in decision making about sharing of data and images are discussed.

Key Points

- *Regulated availability of patient-level data of published clinical studies (data-sharing) is expected.*
- *Expected benefits include verification/advancement of knowledge, reduced cost/time of research, clinical improvement.*
- *Potential drawbacks include faults in patients' identity protection and data misinterpretation.*

Introduction

In clinical research, spontaneous *data sharing* is not yet as common as it is in other fields such as genetics, astronomy or physics [1]. However, the concept of data sharing has been suggested for many reasons, including the patient-centred nature of medical research and healthcare and the expectation

✉ Francesco Sardanelli
   francesco.sardanelli@unimi.it

1 Department of Biomedical Sciences for Health, Università degli Studi di Milano, Via L. Mangiagalli 31, 20133 Milan, Italy

2 Radiology Unit, IRCCS Policlinico San Donato, Via Morandi 30, 20097 San Donato Milanese Milan, Italy

3 PhD Course in Integrative Biomedical Research, Università degli Studi di Milano, Via L. Mangiagalli 31, 20133 Milan, Italy

4 Departments of Radiology and Epidemiology, Erasmus University Medical Center, PO Box 2040, Rotterdam, The Netherlands

5 Department of Health Policy and Medicine, Harvard School of Public Health, Harvard University, 677 Huntington Avenue, Boston, MA 02115, USA

6 Screening and Test Evaluation Program, School of Public Health, Sydney Medical School, University of Sydney, Edward Ford Building, Room A27, Sydney, NSW 2006, Australia

7 Unit of Diagnostic and Interventional Radiology, IRCCS Istituto Ortopedico Galeazzi, Via Riccardo Galeazzi 4, 20161 Milan, Italy

that knowledge from existing data should be maximized to benefit all stakeholders.

Although a transition to data sharing is a process that will take time and planning, those who adopt the principles and practices of open science will likely benefit from it [2, 3]. In addition, the emergence of data sharing as a potential requirement by some agencies and journals warrants attention by the imaging community. Indeed, from July 1st, 2018 the *International Committee of Medical Journal Editors* (ICMJE) will require a data sharing statement as a condition of consideration for publication of clinical trials [4].

In this article, we discuss potential advantages and disadvantages of data sharing.

From *open-access* to *data sharing*

A trend towards larger accessibility to scientific medical knowledge is already visible in the progressive tendency of medical journals in ensuring the *open-access* option, in which the authors or their institutions pay an article-level fee to guarantee the immediate free availability of their papers [5].

In Table 1 we report the policies of all the 18 general imaging journals on access and data sharing [6–17]. This was derived from the current Thomson Reuters list – Radiology, Nuclear Medicine, and Medical Imaging. For comparison, the 17 most-impacted general medicine journals were selected from the current Thomson Reuters list – Medicine, General and Internal [18–35]. Among the 18 imaging journals, four are open access, 12 offer open access as an option (*Radiology* provides free access 12 months after publication), and two do not offer an open-access option. Among the 17 medical journals, six are open access (*The Medical Journal of Australia* only for research articles and case reports), eight offer open access as an option (*Journal of the American Medical Association* [JAMA] provides free access 6 months after publication), two do not offer an open access option, and one (*The New England Journal of Medicine* [NEJM]) provides free access to research articles 6 months after publication. Thus, the open access option is currently widely adopted by both general imaging journals (11/18) and general medicine journals (8/17).

The practice of data sharing entails much more than open access. It is the *regulated availability of the original participant-by-participant data obtained during a study, which may include data not yet analysed.* Among the 18 general imaging journals, data sharing is not even mentioned by 12 journals, encouraged by three, mandatory only upon request in two, and requested by one. Among the 17 general medicine journals, it is not mentioned by seven journals, encouraged by six, requested by three (NEJM only for data obtained by microarray), and considered mandatory only upon request by one (Table 1). In practice, data repository or sharing is currently not

mentioned in the instructions for authors of the majority of general imaging journals (14/18) and major general medicine journals (10/17). Despite individual journals do not mention any policy on data sharing, some publishers such as Elsevier have their own general suggestions, which refer to Open Access [8], even though not immediately visible to the authors when they submit a manuscript. When data sharing is encouraged, authors are informed they should be prepared to provide original study data if requested by the editors.

In recent years, several funding bodies declared the necessity for data sharing. In 2015, the U.S. National Institutes of Health (NIH) expressed its intention to request making the digital data from NIH-funded studies publicly available [36]. Regulatory agencies, specifically the European Medicines Agency, have requested greater data sharing by companies manufacturing drugs and clinical devices. Influential organizations such as the World Health Organization and the U.S. National Academy of Medicine published reports asking for responsible sharing of data from clinical trials [37]. Also, several foundations, for instance the Alfred P. Sloan Foundation [38], the Bill and Melinda Gates Foundation [39], the Ford Foundation [40], the Gordon and Betty Moore Foundation [41], and the National Science Foundation [42], require data sharing and data management plans for all research grant proposals.

The pharmaceutical industry also plays a role in promoting data sharing. The *Yale University Open Data Access* (YODA) project [43] performs independent scientific review of investigators' requests for pharmaceutical and medical data from clinical trials on devices marketed by Johnson & Johnson, including both full clinical study reports and participant-level data. Notably, the YODA project has obtained permission to make independent decisions about the release of Johnson & Johnson's clinical trial data. This project establishes a process in which requests are judged fairly and decisions are made by an independent academic partner, a model that could be applied to other fields of medicine [44].

Another example is the *Academic Research Organization Consortium for Continuing Evaluation of Scientific Studies – Cardiovascular* (ACCESS CV) [45]. They propose a secure method for sharing patient-sensitive data that combines the protection of patients' identity with the legitimate desire of the scientific community for data access and the viewpoint of the researchers who created the database. This approach consists of the following steps: (1) After publication of the primary results of a trial, researchers interested in the study data may send a request to the trial's publication committee; (2) Twenty-four months after the publication of the primary study, requests should be considered by a review group composed of members of ACCESS CV not involved in the trial, the trial principal investigator, a trial statistician, and a member of the data and safety monitoring board. This committee evaluates all proposals to approve those that are feasible,

Table 1 Policies on access and data repository or sharing by major general imaging journals and major general medicine journals

Journal	Access[1]	Data repository or sharing[2]
General imaging journals		
Acad Radiol	Open access option	Not mentioned
Acta Radiol	Open access option	Requested
Am J Roentgenol	No open access option	Not mentioned
BMC Med Imaging	Open access	Encouraged
Br J Radiol	Open access option. Articles freely available more than 12 months after publication	Encouraged
Clin Radiol	Open access option	Not mentioned
Eur J Radiol	Open access option	Not mentioned
Eur Radiol	Open access option	Not mentioned
Invest Radiol	Open access option	Not mentioned
Iran J Radiol	Open access	Not mentioned
J Am Coll Radiol	Open access option	Not mentioned
JBR-BTR	Open access	Not mentioned
Jpn J Radiol	No open access option	Upon request
Korean J Radiol	Open access	Not mentioned
Radiol Med	Open access option	Encouraged
Radiologe	Open access option	Not mentioned
Radiology	Open access option. Articles freely available 12 months after publication	Upon request
Rofo	Open access option	Not mentioned
General medicine journals		
Am J Med	Open access option	Not mentioned
Ann Intern Med	No open access option	Encouraged
BMC Medicine	Open access	Encouraged
Br J Gen Pract	Open access option	Not mentioned
Br Med Bull	Open access option	Not mentioned
BMJ	Open access option	Encouraged
BMJ Open	Open access	Encouraged
CMAJ	Open access option	Requested only for clinical trials of drugs and medical devices
Dtsch Arztebl Int	Open access	Not mentioned
Eur J Clin Invest	Open access option	Encouraged
Int J Med Sci	Open access	Not mentioned
Lancet	Open access option	Encouraged
JAMA	Open access option. Research articles freely available 6 months after publication	Upon request
Med Clin North Am[3]	No open access option	Not mentioned
Med J Aust	Open access for research articles and case reports	Not mentioned
New England J Med	Original articles and special articles freely available 6 months after publication.	Requested for data obtained by microarray
PLoS Med	Open access	Requested

Note: Imaging journals were selected for being general (not subspecialty) journals from the Thomson Reuters list – Radiology, Nuclear Medicine, and Medical Imaging (n=18). For comparison, the general medicine journals from the first quartile were selected from the Thomson Reuters list – Medicine, General and Internal (n=17)

[1] Most journals offer free accessibility for selected articles

[2] Despite individual journals do not mention any policy on data sharing, some publishers (e.g. Elsevier) have their own general rules to which refer to. Moreover, when data sharing is encouraged, authors are informed to be prepared to provide original study data if requested by the editors

[3] Publishes only invited reviews

hypothesis-based, non-duplicative, and guided by investigators with technical capability and a plan for publication. The period of 24 months is chosen to secure the database and to allow the original investigators to perform their own pre-planned secondary analyses; (3) All requests and subsequent decisions will be posted on an ACCESS CV Web portal, ideally within 60 days [45].

In the field of radiology, data sharing also means accessibility to medical images. Indeed, "Images are more than pictures, they are data" [46]. This implies access to the images produced in a given study for additional reading, interpretation, and extraction. To this end, several image repositories were created. An example is the XNAT Central [47, 48], a publicly accessible data repository based on the XNAT open-source platform which hosts a wide variety of research imaging datasets, especially from neuroimaging, but also from oncology, orthopaedics and cardiology. Other examples are *The Cancer Imaging Archive* [49] and the *Lung Image Database Consortium* [50].

Such repositories may be very helpful in several fields, especially for image biomarker development, radiomics and machine learning, each field demanding different approaches. Moreover, the integration, standardization and analysis of these data poses a big challenge, the solution to which may be addressed using cognitive computing. An example of cognitive computing is the system developed by IBM named Watson (IBM Watson Health Imaging, Armonk, NY, USA). It strives to organize available information and present it in a contextually relevant, probability-driven manner to assist healthcare professionals in an objective manner, whether at a reading workstation or at the point-of-care [51]. An important change is underway. To make datasets from medical research publicly available in a timely fashion requires regulations that maximize the benefits and minimize the risks [52, 53]. Indeed, data sharing provides a potential for stimulating new ideas, avoiding duplication of trials, and enhance transparency [36, 54–57] as well as increasing collaboration and interdisciplinary research [1, 58, 59]. However, at the same time, sharing clinical data presents some risks, burdens and challenges such as the need to preserve the privacy of patients, to defend the legitimate economic interests of the sponsors, and to guard against invalid secondary analyses potentially undermining trust in clinical trials or otherwise harming public health [36, 37, 53, 60].

Potential benefits of data sharing

These can be subdivided into: (i) *verification and advancement in knowledge*; (ii) *reduced cost and time for clinical research*; and (iii) *clinical improvement* (Fig. 1).

Verification and advancement in knowledge

The first potential implication of data-sharing is the *verification by independent authors* of the results presented in a given publication. When data are shared, they may be used by other researchers to perform *alternative or supplementary analyses*. This 'second-hand' analysis may show results in support of the initial findings or could reveal errors or inconsistencies in the original research, or could identify issues needing extended analysis.

In other cases, data sharing can allow elucidation of *new results*. New findings can be disclosed starting from hypotheses not considered by the original study team. New insights can be presented from existing data but not yet analysed in the original publication(s). Also, investigators may be interested in *performing the analysis of datasets coming from various sources to enhance precision*, i.e. to perform reproducibility analyses across different databases, regarding established theories or new hypotheses. In fact, reproducibility analysis is crucial for emergent topics in radiology such as standardization of imaging biomarkers, especially from magnetic resonance imaging [61]. The availability of databases from different studies could allow for this gap to be filled and could help in translating new imaging biomarkers into clinical practice [62]. In this regard, reproducibility analysis could become one of the main advantages of data sharing.

The introduction of *registries of patients affected with a defined disease* could be considered a primitive form of data sharing [63, 64], important not only for widespread diseases, such as cancers, but especially for rare diseases.

Another approach of spontaneous data sharing is that underlying *individual patient data meta-analyses* [65]. Authors of an individual patient data meta-analysis typically contact the authors of each eligible study asking to share their data, with the aim of creating a new unique individual-patient database. Of note, the power of the individual-patient data approach is higher than that of conventional (study-level) meta-analyses, which rely on complex statistical methods [66]. For instance, in a study published by Marinovich et al. [67] on the agreement between MRI and pathological breast tumour size after treatment, a total of 24 studies (1,228 patients) were eligible for inclusion, but only eight of these contributed to the individual-patient data analysis for a total of 300 patients. Had regulated data sharing been in place, that individual-patient data meta-analysis would have included a much richer dataset. Moreover, data sharing could boost a wider adoption of health technology assessment. Indeed, in the context of a new product evaluation, data sharing may be useful in the validation level, requiring a high number of data/images, rather than at the initial development level.

Another potential advantage of data sharing is *to reduce the publication of false studies*, especially when the data are intentionally falsified. Recently, 64 articles were retracted from

Fig. 1 Expected pros and cons of data sharing. *IPD* individual patient data meta-analyses

Verification/advancement in knowledge
- Alternative/supplementary analyses
- New results
- Enhanced multidisciplinary cooperation
- Large-scale results (e.g. IPD meta-analysis)
- Reduced publication of false studies
- Estimation of reproducibility

Reduced cost and time for clinical research
- Prevention of duplication of clinical trials

Clinical improvement

The authors could not exploit the entire potential of the database

Possible failure in patients' privacy protection
- Risk of re-identification

Technical barriers
- Data conformity
- Poor dataset documentation

PROS   CONS

DATA SHARING

ten *Springer* journals after editorial checks found fake email addresses, and subsequent internal investigations uncovered fabricated peer-review reports [68]. This retraction came only a few months after *BioMed Central* had retracted 43 articles for the same reason; however, this phenomenon involved most major publishers such as also SAGE, Elsevier, Informa, and Lippincott Williams & Wilkins [69]. Data sharing might discourage data creation and manipulation, potentially more detectable in a complete database than in reported results.

Reduced cost and time for clinical research

Data sharing could potentially lead to an optimization of time and costs of clinical research by *preventing the duplication of trials* [70, 71]. For example, costs for the stipulation of insurances for patients' coverage, the purchase of materials or the salaries of the staff responsible for data collection can be avoided. In addition, using an existing shared database, the new results could be obtained many years prior to those derived from a new clinical study.

Clinical improvement

An effect in terms of clearer evidence on the safety and effectiveness of diagnostic procedures and therapies, improving public healthcare [72–74], may be considered the final aim of data sharing. To avoid the loss of findings contained in the original dataset and not used for the primary publication(s) could play a role in this direction [53]. Institutions sharing their data could obtain a more comprehensive picture about the benefits and risks of a medical decision. However, a real clinical improvement from data sharing is a hypothesis that still needs to be demonstrated.

Potential drawbacks from data sharing

The sharing of clinical databases raises several concerns (see Fig. 1). One of the reasons not to share data is that researchers are evaluated competitively, based on the quality and number of articles published during their career, so they may worry that other people will use their data and efforts to produce new publications. The potential for secondary analyses contradicting initially reported results may be a deterrent. Authors may not be willing to share data that had cost them great effort and resources. However, reciprocally, they would also directly benefit from using someone else's data.

Bierer et al. [75] recently suggested formalizing 'data authorship' as an incentive to data sharing: "as a matter of fairness and as a matter of providing an incentive for data sharing, the persons who initially gathered the data should receive appropriate and standardized credit that can be used for academic advancement, for grant applications, and in broader situations".

Another concern is the potential for fault in the patient identity protection caused by the transmission of sensitive information. Data must be de-identified: *de-identification*, not simply anonymization, consists of transforming a dataset so that the back identification of individuals becomes impossible or extremely difficult. Different regulations may require different degrees of de-identification, particularly in the absence of informed consents specifying the possibility of data sharing. De-identification can be achieved with different types of data transformations that must ensure patient privacy without affecting data quality [76]. However, the de-identified data do not eliminate all risks of re-identification. Moreover, the reduction of this risk to zero may destroy or significantly impair the utility of the data for subsequent analysis or

verification. For these reasons, the stipulation of Data Use Agreements (DUAs) is considered a useful strategy and best practice for increasing the benefits and mitigating the risks of clinical data sharing [77]. Specifically, DUAs address important issues such as limitations on date usage, obligations to data safeguard, liability for harm arising from data usage and publication, and privacy rights that are associated with transfer of confidential or protected data. In contrast, the U.S. Office for Human Research Protections stated that there is no need for separate consent from trial participants for the sharing of de-identified data [4].

A limitation to the adoption of data sharing can originate from *technical barriers*. The image conformity is influenced by vendor, modality, and acquisition parameters on the one hand; and by image post-processing manufacturer, reconstruction parameters, and software versions, on the other hand. An example is represented by the use in magnetic resonance of arbitrary units that clearly depend on the specific vendor and model, making a between-study comparison impossible. A way to overcome this limitation could be a drastic standardization, with manufacturers defining new shared standards.

Another intrinsic barrier to data sharing could be the *poor documentation of datasets*, especially if not documented in English. Moreover, important information about methodology might not be contained immediately in the database or immediately retrievable. All these issues should be considered when planning for potential data sharing of research.

To share or not to share?

In conclusion, in a world that moves towards greater transparency and privacy protection, data sharing stands between these two competing interests. Not all concerns on data sharing have already been solved and many questions remain to be addressed: Who is the rightful owner of the data? What is the role of individual patients and advocacy groups in decision making about sharing of data and images? Should Ethics Committees change their approach for study approval? And how? What is the exact role of institutions, especially public ones, that funded the original study? Should patient advocacy groups and funding organizations be involved in decision making about data sharing? These issues must be regulated.

Despite all the above-described issues relating to data sharing, a transition to a more open medical science has begun. If benefits of data sharing will be more and more perceived as prevailing over harms therefrom, this option will win. Researchers and institutions who first seize this opportunity will be on the wave-front of an innovation likely to be in favour of patients and public health. Radiologists should be kept informed of this emerging issue. It is time to share!

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is prof. Francesco Sardanelli.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent This study is not on human subjects

Ethical approval Institutional Review Board approval was not required because this is an editorial article.

Methodology This is an editorial article

References

1. Ross JS, Lehman R, Gross CP (2012) The importance of clinical trial data sharing: toward more open science. Circ Cardiovasc Qual Outcomes 5:238–240
2. Boulton G, Rawlins M, Vallance P, Walport M (2011) Science as a public enterprise: the case for open data. Lancet 377:1633–1635
3. Walport M, Brest P (2011) Sharing research data to improve public health. Lancet (London, England) 377:537–539
4. Taichman DB, Sahni P, Pinborg A et al (2017) Data sharing statements for clinical trials—A requirement of the International Committee of Medical Journal Editors. N Engl J Med 376:2277–2279
5. Sconfienza LM, Sardanelli F (2013) Radiological journals in the online world: should we think open? Eur Radiol 23:1175–1177
6. RSNA open access policy. Radiological Society of North America web site. http://pubs.rsna.org/page/openaccess. Accessed 29 July 2017
7. Publish open access with Springer. Springer web site. http://www.springer.com/de/open-access. Accessed 29 July 2017
8. Open access. Elsevier web site. https://www.elsevier.com/about/open-science/open-access. Accessed 29 July 2017
9. Online submission and review system. Investigative Radiology web site. http://edmgr.ovid.com/ir/accounts/ifauth.htm. Accessed 29 July 2017
10. American Journal of Roentgenology web site. http://www.ajronline.org/. Accessed 29 July 2017
11. Acta Radiologica open submission guidelines. SAGE Publishing web site. https://us.sagepub.com/en-us/nam/acta-radiologica-open/journal202176#description. Accessed 29 July 2017
12. Open access policy. The British Institute of Radiology web site. http://www.birpublications.org/page/oapolicy. Accessed 29 Jul 2017

13. Guidelines for authors. Rofo-Fortschr Rontg web site. http://roefo.thieme.de/documents/10157/18614/RoeFo-Autorenhinweise_Englisch-2017.pdf/ef85bdcc-03d3-41d4-8088-215c16528db9. Accessed 29 July 2017

14. BioMed Central Medical Imaging web site. https://bmcmedimaging.biomedcentral.com/about. Accessed 29 July 2017

15. Publication instructions for authors. Korean Journal of Radiology web site. https://www.kjronline.org/index.php?body=Instruction. Accessed 29 July 2017

16. Open access statement. Iranian Journal of Radiology web site. http://iranjradiol.com/?page=public_pages&name=Open Access Statement. Accessed 29 July 2017

17. Journal of the Belgian Society of Radiology web site. http://www.jbsr.be/about/. Accessed 29 July 2017

18. The New England Journal of Medicine web site. http://www.nejm.org/page/about-nejm/history-and-mission. Accessed 29 July 2017

19. Information for authors. The Lancet web site. http://thelancet.com/lancet/information-for-authors/open-access. Accessed 29 July 2017

20. Instruction for authors. Journal of the American Medical Association web site. http://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecPublicAccess. Accessed 29 July 2017

21. Information for authors. Annals of Internal Medicine web site. http://annals.org/aim/pages/authors. Accessed 29 July 2017

22. Resources for authors. British Medical Journal web site. http://www.bmj.com/about-bmj/resources-authors. Accessed 29 July 2017

23. Why publish with PLOS Medicine? PLoS Medicine web site. http://journals.plos.org/plosmedicine/s/why-publish-with-plos-medicine. Accessed 29 July 2017

24. Fees and funding. BioMed Central Medicine web site. http://bmcmedicine.biomedcentral.com/submission-guidelines/fees-and-funding. Accessed 29 July 2017

25. The American Journal of Medicine open access option. Elsevier web site. https://www.elsevier.com/journals/the-american-journal-of-medicine/0002-9343/open-access-options. Accessed 29 July 2017

26. CMAJ Open. Canadian Medical Association Journal Open web site. http://cmajopen.ca/site/misc/about.xhtml. Accessed 29 July 2017

27. Deutsches Arzteblatt International web site. https://www.aerzteblatt.de/int/about-us. Accessed 29 July 2017

28. MJA Open. Medical Journal of Australia web site. https://www.mja.com.au/open. Accessed 29 July 2017

29. Open access. Oxford Academic web site. https://academic.oup.com/journals/pages/open_access. Accessed 29 July 2017

30. BJGP editorial process & policies. British Journal of General Practice web site. http://bjgp.org/authors/bjgp-editorial-process-and-policies. Accessed 29 July 2017

31. OnlineOpen. Wiley Author Services web site. https://authorservices.wiley.com/author-resources/Journal-Authors/licensing-and-open-access/open-access/onlineopen.html. Accessed 29 July 2017

32. BioMed Central web site. https://www.biomedcentral.com/about. Accessed 29 July 2017

33. British Medical Journal Open web site. http://bmjopen.bmj.com/pages/about/. Accessed 29 July 2017

34. Medical Clinics of North America open access option. Elsevier web site. https://www.elsevier.com/journals/medical-clinics-of-north-america/0025-7125/open-access-options. Accessed 29 July 2017

35. Instruction for authors. International Journal of Medical Sciences web site. http://www.medsci.org/ms/author. Accessed 29 July 2017

36. Collins FS, Tabak LA (2014) Policy: NIH plans to enhance reproducibility. Nature 505:612–613

37. Sharing clinical trial data: maximizing benefits, minimizing risk - PubMed - NCBI. https://www.ncbi.nlm.nih.gov/pubmed/25590113. Accessed 18 Jan 2017

38. Grants. Alfred P. Sloan Foundation web site. https://sloan.org/grants/apply#tab-grant-proposal-guidelines/. Accessed 1 April 2017

39. Open access policy. Bill and Melinda Gates Foundation web site. http://www.gatesfoundation.org/how-we-work/general-information/open-access-policy. Accessed 1 April 2017

40. Ford Foundation expands Creative Commons licensing for all grant-funded projects. Ford Foundation web site. https://www.fordfoundation.org/the-latest/news/ford-foundation-expands-creative-commons-licensing-for-all-grant-funded-projects/. Accessed 1 April 2017

41. Data sharing philosophy. Gordon and Betty Moore Foundation web site. https://www.moore.org/docs/default-source/Grantee-Resources/data-sharing-philosophy.pdf. Accessed 1 April 2017

42. Dissemination and sharing of research results. National Science Foundation web site. https://www.nsf.gov/bfa/dias/policy/dmp.jsp. Accessed 1 April 2017

43. Krumholz HM, Waldstreicher J (2016) The Yale Open Data Access (YODA) project — A mechanism for data sharing. N Engl J Med 375:403–405

44. Krumholz HM, Ross JS (2011) A model for dissemination and independent analysis of industry data. JAMA 306:1593–1594

45. Academic Research Organization Consortium for Continuing Evaluation of Scientific Studies–Cardiovascular (ACCESS CV), Patel MR, Armstrong PW, Bhatt DL et al (2016) Sharing data from cardiovascular clinical trials—A Proposal. N Engl J Med 375:407–409

46. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577

47. Herrick R, Horton W, Olsen T et al (2016) NeuroImage XNAT central: open sourcing imaging research data. NeuroImage 124:1093–1096

48. XNAT web site. https://www.xnat.org/about/. Accessed 13 April 2017

49. Clark K, Vendt B, Smith K et al (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26:1045–1057

50. Armato SG, McLennan G, Bidaut L et al (2011) The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38:915–931

51. Chen Y, Elenee Argentinis JD, Weber G (2016) IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research. Clin Ther 38:688–701

52. Loder E (2013) Sharing data from clinical trials: where we are and what lies ahead. BMJ 347:f4794–f4794

53. Mello MM, Francer JK, Wilenzick M et al (2013) Preparing for responsible sharing of clinical trial data. N Engl J Med 369:1651–1658

54. Anderson BJ, Merry AF (2009) Data sharing for pharmacokinetic studies. Paediatr Anaesth 19:1005–1010

55. Gøtzsche PC (2011) Why we need easy access to all data from all clinical trials and how to accomplish it. Trials 12:249

56. Berlin JA, Morris S, Rockhold F et al (2014) Bumps and bridges on the road to responsible sharing of clinical trial data. Clin Trials 11:7–12

57. Peat G, Riley RD, Croft P et al (2014) Improving the transparency of prognosis research: the role of reporting, data sharing, registration, and protocols. PLoS Med 11:e1001671

58. Milia N, Congiu A, Anagnostou P et al (2012) Mine, yours, ours? Sharing data on human genetic variation. PLoS One 7:e37552

59. Lee ES, McDonald DW, Anderson N, Tarczy-Hornoch P (2009) Incorporating collaboratory concepts into informatics in support of

translational interdisciplinary biomedical research. Int J Med Inform 78:10–21

60. Antman E (2014) Data sharing in research: benefits and risks for clinicians. BMJ 348:g237

61. Sardanelli F (2017) Trends in radiology and experimental research. Eur Radiol Exp https://doi.org/10.1186/s41747-017-0006-5

62. Golay X (2017) The long and winding road to translation for imaging biomarker development: the case for arterial spin labelling (ASL). Eur Radiol Exp 3. https://doi.org/10.1186/s41747-017-0004-7

63. Grill JD, Holbrook A, Pierce A et al (2017) Attitudes toward potential participant registries. J Alzheimers Dis 56:939–946

64. Kasenda B, von Elm E, You J et al (2014) Prevalence, characteristics, and publication of discontinued randomized trials. JAMA 311:1045–1051

65. Clarke MJ, Stewart LA (1997) Meta-analyses using individual patient data. J Eval Clin Pract 3:207–212

66. Phi X-A, Houssami N, Obdeijn I-M et al (2015) Magnetic resonance imaging improves breast screening sensitivity in BRCA mutation carriers age ≥50 years: evidence from an individual patient data meta-analysis. J Clin Oncol 33:349–356

67. Marinovich ML, Macaskill P, Irwig L et al (2015) Agreement between MRI and pathologic breast tumor size after neoadjuvant chemotherapy, and comparison with alternative tests: individual patient data meta-analysis. BMC Cancer 15:662

68. Retraction of articles from Springer journals. http://www.springer.com/gp/about-springer/media/statements/retraction-of-articles-from-springer-journals/735218. Accessed 12 March 2017

69. Qi X, Deng H, Guo X (2017) Characteristics of retractions related to faked peer reviews: an overview. Postgrad Med J 93:499–503

70. Rathi V, Dzara K, Gross CP et al (2012) Sharing of clinical trial data among trialists: a cross sectional survey. BMJ 345:e7570

71. Zarin DA (2013) Participant-level data and the new frontier in trial transparency. N Engl J Med 369:468–469

72. Farrar JT, Troxel AB, Haynes K et al (2014) Effect of variability in the 7-day baseline pain diary on the assay sensitivity of neuropathic pain randomized clinical trials: An ACTTION study. Pain 155:1622–1631

73. Gabler NB, French B, Strom BL et al (2012) Validation of 6-minute walk distance as a surrogate end point in pulmonary arterial hypertension trials. Circulation 126:349–356

74. Gabler NB, French B, Strom BL et al (2012) Race and sex differences in response to endothelin receptor antagonists for pulmonary arterial hypertension. Chest 141:20–26

75. Bierer BE, Crosas M, Pierce HH (2017) Data authorship as an incentive to data sharing. N Engl J Med 376:1684–1687

76. Prasser F, Bild R, Kuhn KA (2016) A generic method for assessing the quality of de-identified health data. Stud Health Technol Inform 228:312–316

77. Barocas S, Nissenbaum H (2014) Big data's end run around anonymity and consent. In: Lane J, Stodden V, Bender S, Nissenbaum H (eds) Privacy, big data, and the public good. Cambridge University Press, New York, pp 44–75