



CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia



Methodological letter

Transparency of my results. How to deposit my study data in an open access repository?

Transparencia de mis resultados. ¿Cómo depositar los datos de mi estudio en un repositorio de acceso abierto?

Raul Muñoz Garcia,^a Guillem Cebrián Grifol^{b,*}

^a Unidad de Soporte a la Investigación, Institut d'Investigació i Innovació Parc Taulí (I3PT-CERCA), Barcelona, Spain

^b Unidad de Gestión del Conocimiento, Institut d'Investigació i Innovació Parc Taulí (I3PT-CERCA), Barcelona, Spain

Introduction

Science has significantly changed in recent decades, marked by a paradigm shift that has reconfigured its practice and purpose, promoting greater transparency, collaboration and accessibility to knowledge. This movement promotes the free spread of knowledge and stresses the importance of sharing data and results of scientific activity in an open and collaborative manner, thereby accelerating advances in science. Throughout these years research personnel in the field of surgery have been included in these transformations and practices, and have witnessed changes, including those in publication models and research result dissemination.

Within the framework of open science, Spanish and European regulations have also added certain requirements regarding the publication of research results. For projects financed with public funds, the deposit of a copy of the final accepted version for scientific publications is requested, as well as the associated data in open access repositories.¹ This change in regulations has led funding agencies and publishers to incorporate these requirements in the open tenders and review criteria, respectively. As a result, researchers may ask themselves: What is an open access repository? What does

depositing my data in one of these entail and what are the benefits? What aspects should I consider when depositing? Or who can I contact for guidance within my organisation? This article aims to briefly explain the necessary steps to follow for depositing in a repository and thus provide answers to the most common doubts.

What is an open access repository?

An open access repository is a digital platform where different types of academic-scientific content, mainly research results, are openly stored and made available to the community. These repositories play a crucial role in promoting open access, research transparency, and the preservation and dissemination of scientific knowledge.² They also help meet the requirements of funding agencies and journals. However, depositing is not synonymous with publishing the data freely and openly. Certain formulas allow access to published data to be restricted to a greater or lesser extent (use licenses, data embargo, etc.). Every digital object deposited in a repository will be assigned a DOI (Digital Object Identifier) to provide unique, permanent identification, facilitating its citation and monitoring over time.

* Corresponding author.

E-mail address: gcebrian@tauli.cat (G. Cebrián Grifol).

<https://doi.org/10.1016/j.cireng.2024.07.006>

2173-5077/© 2024 AEC. Published by Elsevier España, S.L.U. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

How do I deposit my data?

Which repository should I choose?

Before publishing a dataset, you must first choose the appropriate repository. There are different types of repositories (editorial or institutional, thematic or multidisciplinary, etc.), so certain aspects must be considered during the choice:

- Is my data easily adaptable to the required format?
- Does the volume of the files to be deposited meet my needs?
- Is data restriction allowed if necessary?
- Are the licenses offered in keeping with my interests?
- Is a specific DOI generated for my data?
- Are there costs associated with its use?
- Is there a data preservation policy?
- Is there version control of the datasets?
- Does it comply with the requirements of European aid or the standards of journals or the scientific community?

It is strongly recommended to choose repositories that enable data depositing in accordance with the FAIR principles (Findable, Accessible, Interoperable, Reusable)³ and under the guidelines of the European Open Science Cloud (EOSC), since they are the ones that guarantee compliance of ethical and legal standards. Re3data.org (<https://www.re3data.org/>), is a highly useful tool for choosing a repository, as it compiles metadata from numerous specialised research repositories.

How should the dataset be prepared for depositing?

The next step is preparing the data for the repository, according to the guidelines of the chosen repository. The most common requirements include the **anonymisation or pseudonymisation of the data** (if it contains identifying data), the **saving of these in open format**, sustainable and non-proprietary files to facilitate their long-term preservation, and the description of the data. It is essential to explain how the data has been created, its context, structure and content to facilitate its interpretation and reuse. To do this when depositing the following is required: **description of the metadata**, which is the data that describes the dataset (title,

author and affiliation, contact, language, etc.); **Readme file**, a mandatory document in .txt format that allows you to communicate important information about the dataset (title, methods, summary of data and files, reuse conditions, etc.); and, finally, the **codebook or list of variables**, which is usually tabular and provides complete and self-explanatory information for each of the variables in the dataset.

What type of licence should I use?

Before making the deposit, it is important to define the license to use the dataset. The License Chooser tool (<https://chooser-beta.creativecommons.org/>) offered by Creative Commons is very useful for identifying the most appropriate license for each case. Some repositories suggest using the most open licenses possible (CC0 or CC BY) to ensure the greatest reuse of data. It is important to find a balance between openness and data protection, following the principle of “as open as possible, as closed as necessary”.

Who should I approach for dataset review?

Once the repository has been chosen, the data and attached documents have been prepared, and the appropriate license has been assigned, the next step is to curate the dataset. It is recommended that this be reviewed by the research centre’s support staff, as they are familiar with the standards and requirements of the data repository and ensure compliance with open science guidelines. Once the previous points have been completed and this internal review has been carried out, the dataset is prepared for publication (Fig. 1).

Real example of dataset deposit

To improve the understanding of the process of depositing a dataset and the required documentation, a real case from the Institut d’Investigació i Innovació Parc Taulí (I3PT) is shown. Dr. Xavier Serra-Aracil contacted the support staff to be able to deposit a dataset from the TAUTEM project.^{4,5} After several meetings where the characteristics of the dataset were discussed and reviewed, the result of depositing was as follows: Serra-Aracil X. A randomised multicentre non-inferiority phase III trial of chemo-radiotherapy and local

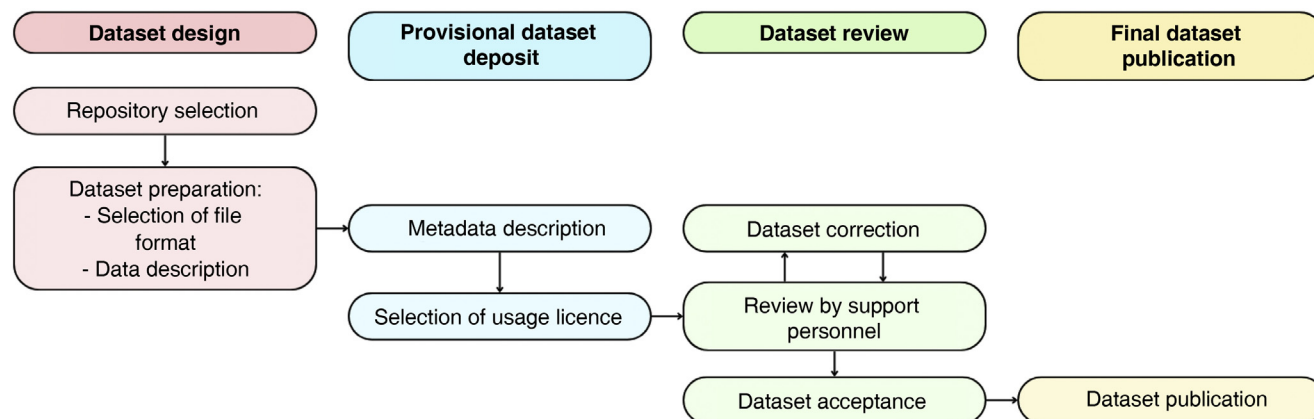


Fig. 1 – Steps to follow for the deposit of a dataset in an open access repository.

excision versus total mesorectal excision in T2–T3ab,N0,M0 rectal cancer (the TAUTEM study). CORA.Repositori de Dades de Recerca. 2024. <https://doi.org/10.34810/data1181>.⁶

Conclusions

Depositing data in open access repositories is essential for modern science, promoting transparency, collaboration and reuse of information. Although there may be initial concerns, recent studies show that researchers find the process easier than expected. When done properly, not only is it risk-free, but open data repository provides considerable benefits.^{7,8} It should be remembered that institutions have specialized support personnel to help in the process, guaranteeing compliance with legal and ethical requirements, which are crucial to avoid errors that compromise the integrity, reproducibility and security of the data. In addition, there are specific infrastructures and tools that facilitate this procedure, making it more accessible to research staff.

Promoting data repository is an ethical responsibility and an opportunity to advance scientific knowledge in a more effective and collaborative way. By taking advantage of available resources and asking for help when necessary, we can all make significant contributions to open science, stepping up scientific progress and enabling broader access to knowledge.

REFERENCES

- Jefatura del Estado. Ley 17/2022, de 5 de septiembre, por la que se modifica la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación [Internet]. Sec. 1, Ley 17/2022 sep 6, 2022 p. 123852-922 [Accessed 30 May 2024]. Available from: <https://www.boe.es/eli/es/l/2022/09/05/17>.
- Melero R, Hernández-San-Miguel J. Acceso abierto a los datos de investigación, una vía hacia la colaboración científica. *Rev Esp Doc Científica*. 2014;37. <http://dx.doi.org/10.3989/redc.2014.4.1154>. e066-e066.
- Wilkinson MD, Dumontier M, Aalbersberg JJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018. <http://dx.doi.org/10.1038/sdata.2016.18>.
- Serra-Aracil X, Pericay C, Golda T, Mora L, Targarona E, Delgado S, et al. Non-inferiority multicenter prospective randomized controlled study of rectal cancer T2-T3s (superficial) N0, M0 undergoing neoadjuvant treatment and local excision (TEM) vs total mesorectal excision (TME). *Int J Colorectal Dis*. 2018;33:241–9. <http://dx.doi.org/10.1007/s00384-017-2942-1>.
- Serra-Aracil X, Pericay C, Badia-Closa J, Golda T, Biondo S, Hernández P, et al. Short-term outcomes of chemoradiotherapy and local excision versus total mesorectal excision in T2-T3ab,N0,M0 rectal cancer: a multicentre randomised, controlled, phase III trial (the TAUTEM study). *Ann Oncol Off J Eur Soc Med Oncol*. 2023;34:78–90. <http://dx.doi.org/10.1016/j.annonc.2022.09.160>.
- Serra-Aracil X, Pericay C, Golda T, Mora L, Targarona E, Delgado S, et al. Non-inferiority multicenter prospective randomized controlled study of rectal cancer T2-T3s (superficial) N0 M0 undergoing neoadjuvant treatment and local excision (TEM) vs total mesorectal excision (TME). *Int J Colorectal Dis*. 2018;33:241–9. <http://dx.doi.org/10.1007/s00384-017-2942-1>.
- Serra-Aracil X, Pericay C, Badia-Closa J, Golda T, Biondo S, Hernández P, et al. Short-term outcomes of chemoradiotherapy and local excision versus total mesorectal excision in T2-T3ab,N0 M0 rectal cancer: a multicentre randomised, controlled, phase III trial (the TAUTEM study). *Ann Oncol Off J Eur Soc Med Oncol*. 2023;34:78–90. <http://dx.doi.org/10.1016/j.annonc.2022.09.160>.
- Stieglitz S, Wilms K, Mirbabaie M, Hofeditz L, Brenger B, López A, et al. When are researchers willing to share their data? – Impacts of values and uncertainty on open data in academia. *PLoS ONE*. 2020;15:e0234172. <http://dx.doi.org/10.1371/journal.pone.0234172>.