

ESTUDIO DE CASO

SOBRE *LOS ARTÍCULOS PUBLICADOS EN ACCESO ABIERTO CON DATOS ASOCIADOS* EN PUBMEDCENTRAL EN UN HOSPITAL UNIVERSITARIO (2020-2024)

Blanca San-José-Montano. Biblioteca Ciencias de la Salud Hospital Clínico San Carlos

Rafael García-Carretero. Hospital Universitario de Móstoles

Manuel Espantaleon Agreda. Biblioteca Ciencias de la Salud Hospital Clínico San Carlos

CIENCIA ABIERTA *Declaración de la Unesco (2021)*

Promueve la accesibilidad, reutilización y colaboración en la investigación científica... en beneficio de la ciencia y la sociedad



PubMed Central (PMC) Repositorios que más fomentan la ciencia abierta, especialmente en el campo biomédico

OFRECE:

- Acceso libre a **texto completo de artículos científicos**, gran difusión (PubMed)
- **Preservación digital** a largo plazo

QUE INCLUYE:

- **Trabajos financiados fondos públicos EEUU y de acceso gratuito** (políticas de acceso abierto NIH) .
- **Revistas completas que han establecido acuerdos editoriales** con la National Library of Medicine (NLM) para depositar automáticamente todos (o parte) de sus contenidos. Las revistas tienen que cumplir ciertos criterios editoriales, políticas de revisión por pares, garantizar acceso abierto, etc.

FILTROS: nuevo filtro 2014 denominado **“associated Data”** (“has associated data[filter]”):

- Material suplementario (“has suppdata[filter]”).
- Declaración de datos (forma de obtención, porque no pueden facilitarse, etc.) (“has data avail[filter]”).
- Trabajos con la citación bibliográfica de los datos utilizados en su elaboración (“has data citations[filter]”).

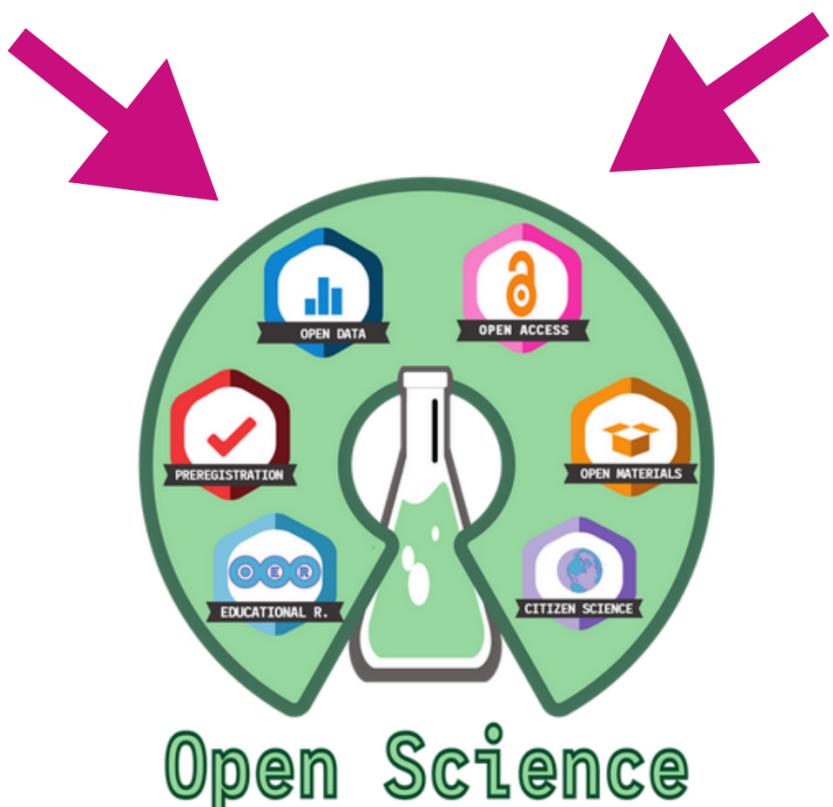
WHY?



PubMed



“Associated Data” solo limitar los resultado de las búsquedas a artículos que tenían enlaces de datos relacionados con el campo (identificadores MEDLINE Databank Sources o el LinkOut (proveedores más recursos).





- Centro de referencia en la formación de profesionales sanitarios de la **Universidad Complutense de Madrid**.
- Ampla trayectoria de actividad investigadora liderada por el **Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC)**.
- **Adaptándose a los nuevos mandatos nacionales y europeos** (*Ley de Ciencia, tecnología e innovación (2022)*, *Ley orgánica del Sistema Universitario (2023)*, *Estrategia Nacional de Ciencia Abierta (2023-27)*. *Plan estatal de investigación científica y técnica y de Innovación (PEICTI, 2024-27)*... **NUEVA EVALUACIÓN CIENTÍFICA**



OBJETIVO

1st

CONOCER COMO REGISTRA LOS ARTÍCULOS PMC CUANDO ESTABLECE QUE TIENE DATOS ASOCIADOS a través de analizar los artículos publicados en acceso abierto incorporados a esta base de datos por autores del hospital Clínico San Carlos durante los últimos 5 años.

2nd

CONOCER LA TIPOLOGÍA DE ESOS DATOS ASOCIADOS, SU TIPO DE DECLARACIÓN Y SI HAN SIDO REGISTRADOS EN ALGÚN REPOSITORIO (tipo de repositorio y cita) a través de la minuciosa revisión de cada uno de sus registros.

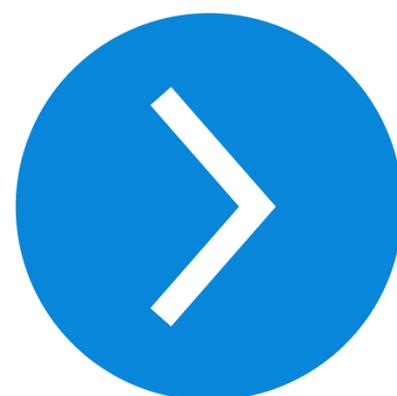
MÉTODO Y RESULTADOS:

Búsqueda en PubMedCentral

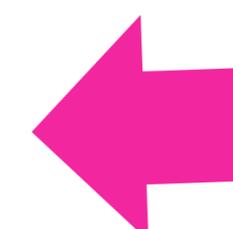
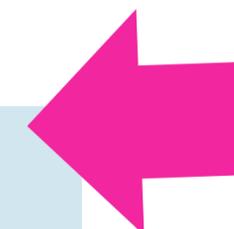
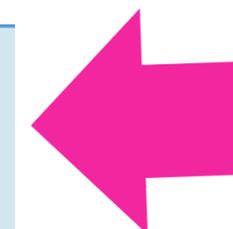
"hospital clinico san carlos"[Affiliation]

CON FILTROS:

- Acceso abierto
- Material Complementario
- Datos asociados
- Declaración de datos
- Citación de Datos
- 2020-2024



TIPO DE FILTRO	ESTRATEGIA	Resultado PMC	RESULTADO REVISIÓN
Publicado en acceso abierto, del año 2020 al 2024	"hospital clinico san carlos"[Affiliation] AND ("open access"[filter] ("2020/01/01"[PubDate] : "2024/12/31"[PubDate]))	2.031	---
Con datos asociados y publicado en acceso abierto, 5 años	"hospital clinico san carlos"[Affiliation] AND ("open access"[filter] AND "has associated data"[Filter]) ("2020/01/01"[PubDate] : "2024/12/31"[PubDate]))	989 (48,7% de 2.031)	965 (eliminado 24 trabajos)
Con datos asociados y Material Complementario asociado	"hospital clinico san carlos"[Affiliation] AND "has associated data"[Filter] AND ("open access"[filter] AND ("2020/01/01"[PubDate] : "2024/12/31"[PubDate]))	556 (56,2% de 989)	521
Datos disponibles o declaración de datos	"hospital clinico san carlos"[Affiliation] AND "has data avail[filter] AND ("open access"[filter] AND ("2020/01/01"[PubDate] : "2024/12/31"[PubDate]))	988 (99,8% de 989)	962 (eliminado 4 sin Declaración datos) quedan 958
Citación de datos	"hospital clinico san carlos"[Affiliation] AND "has data citations[filter] AND ("open access"[filter] AND "has associated data"[Filter]) ("2020/01/01"[PubDate] : "2024/12/31"[PubDate]))	2	1 cita datos creados, otro cita fuente datos terceros



MÉTODO Y RESULTADOS: Análisis cuantitativo resultados - Característica Trabajos recuperados-

Totales artículos publicados

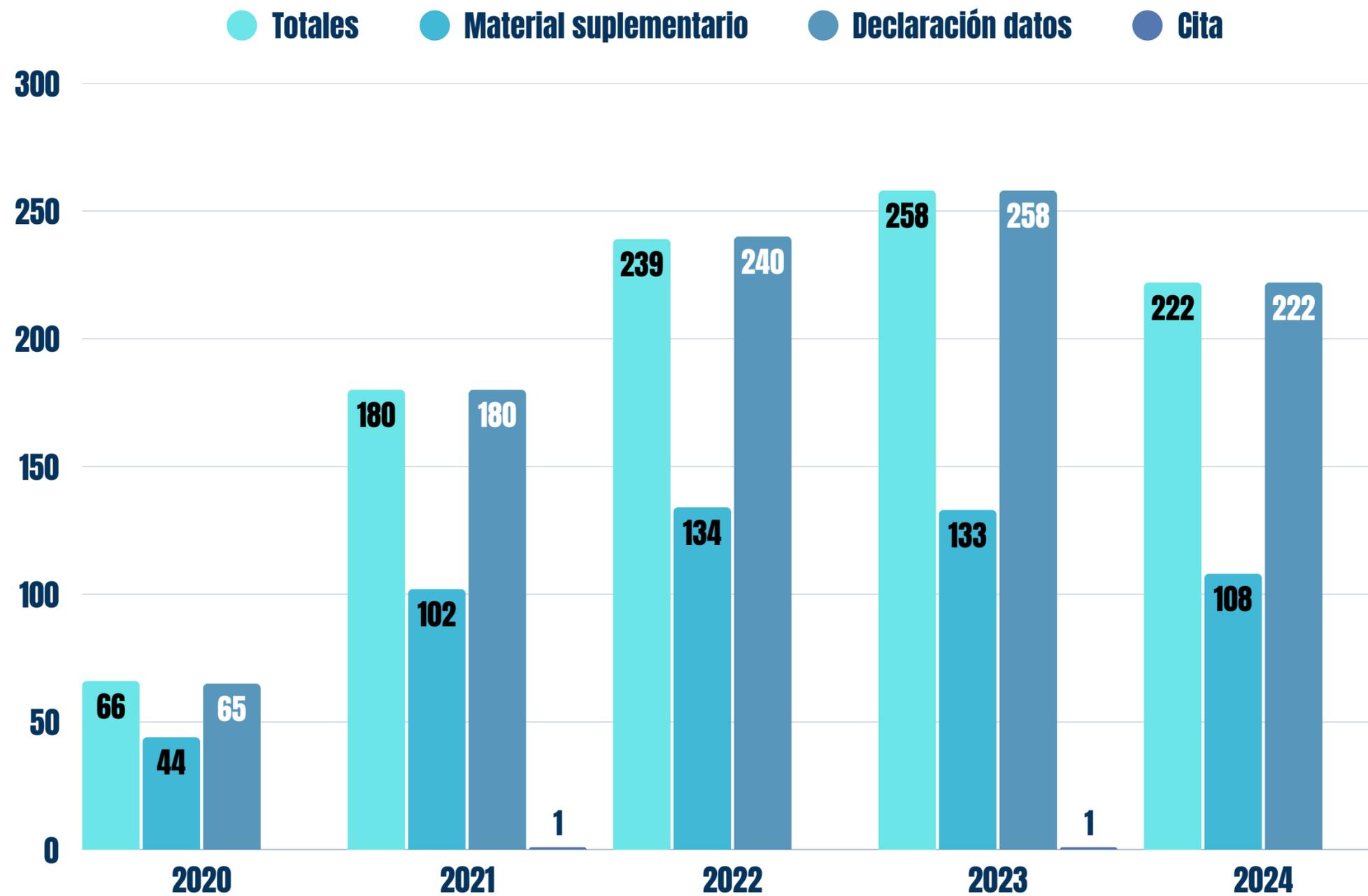
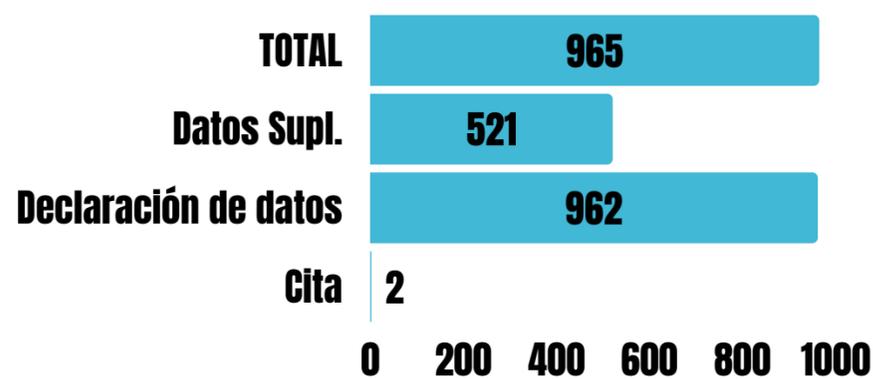


Material Suplementario



Declaración datos

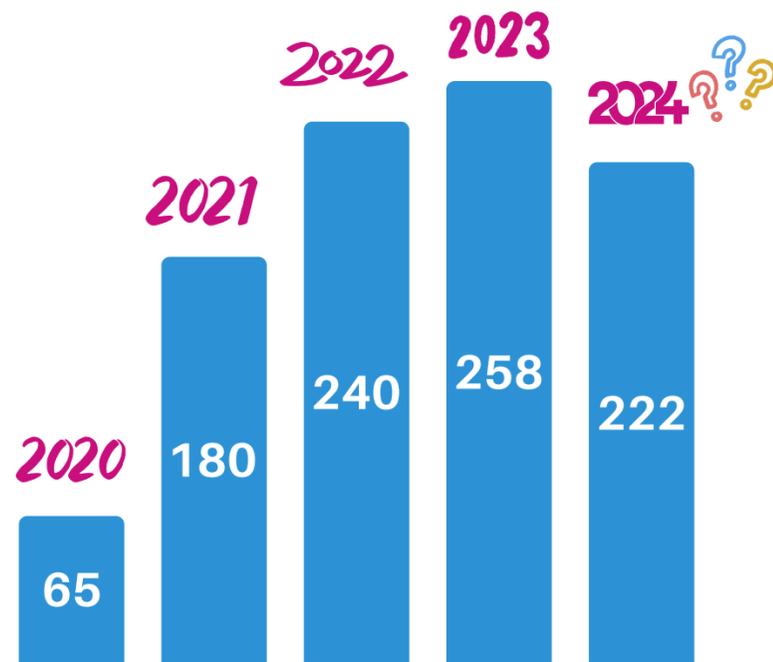
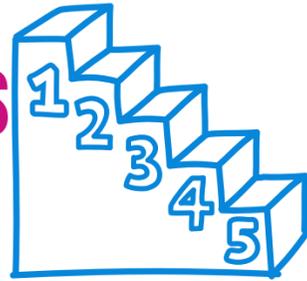
Cita



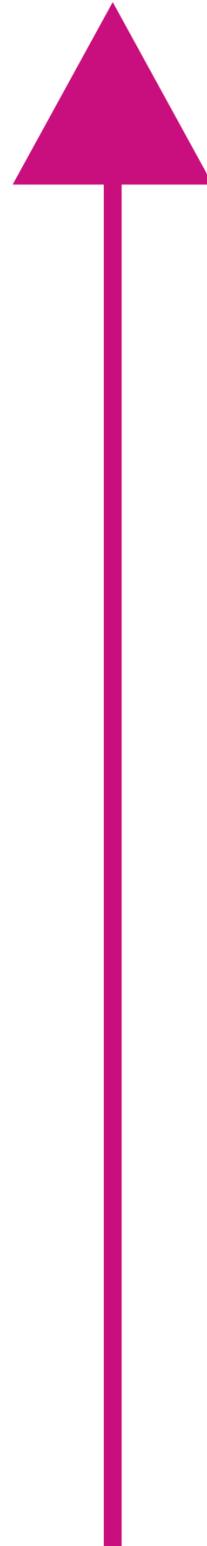
MÉTODO Y RESULTADOS: Análisis cuantitativo resultados - RANKING REVISTAS -

Ranking Revistas

+ 9

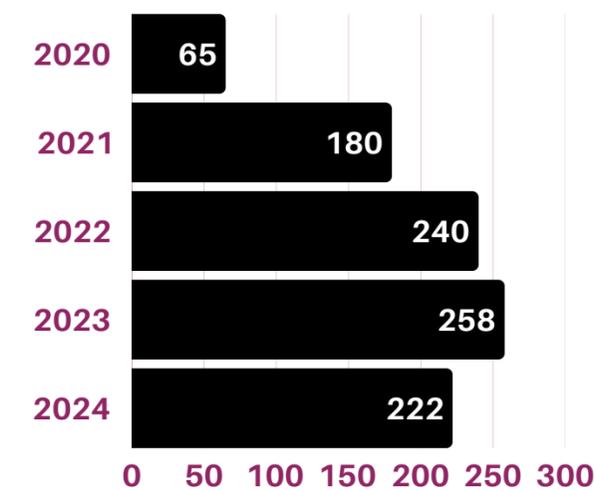
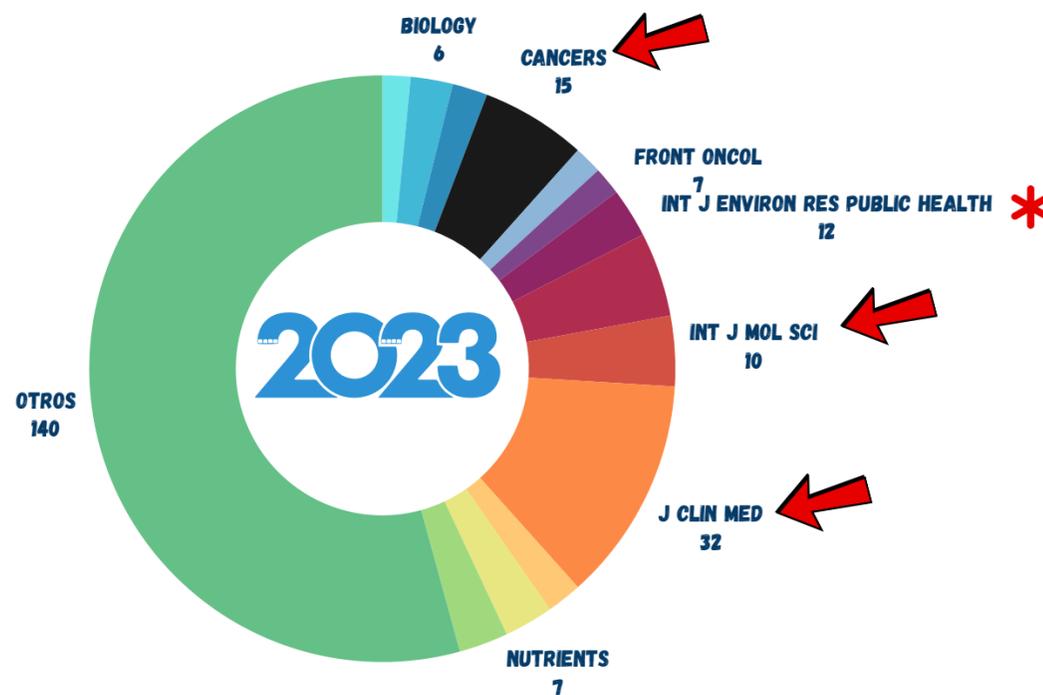
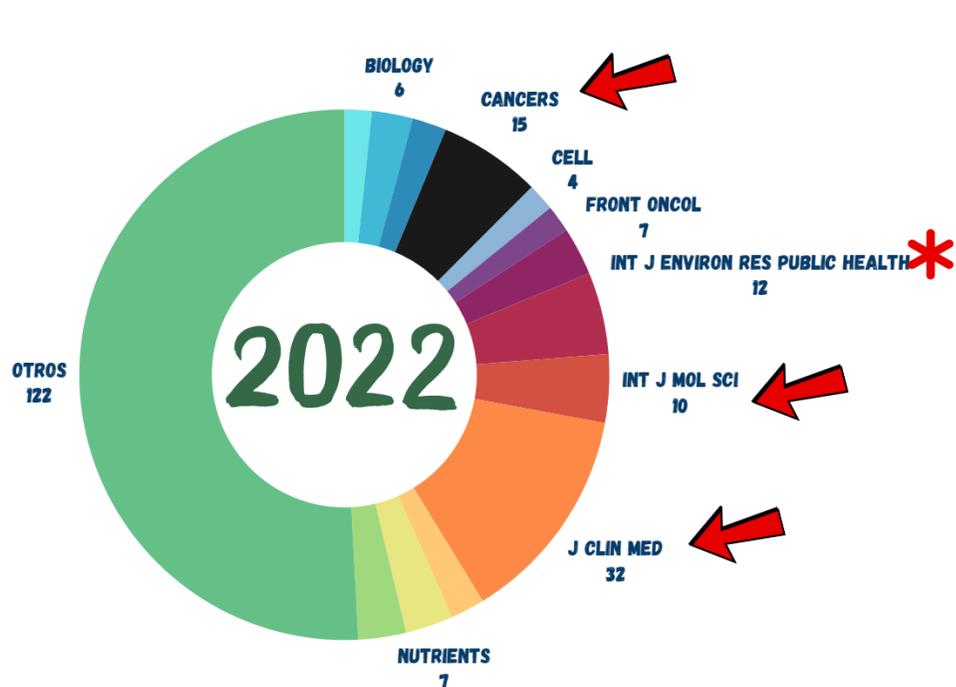
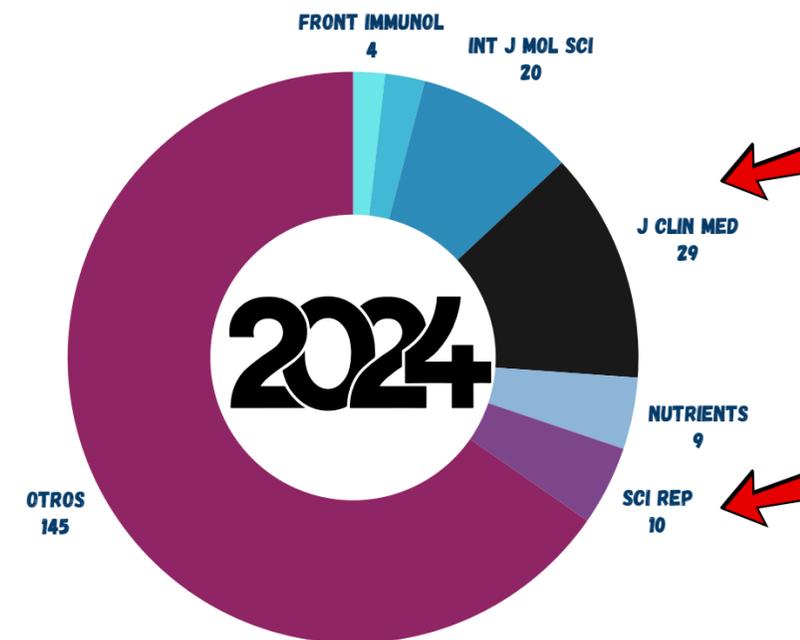
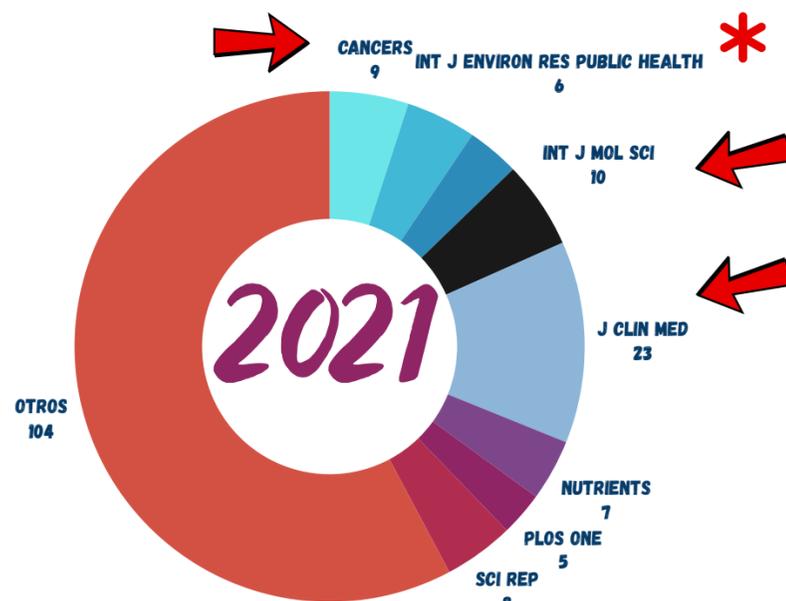
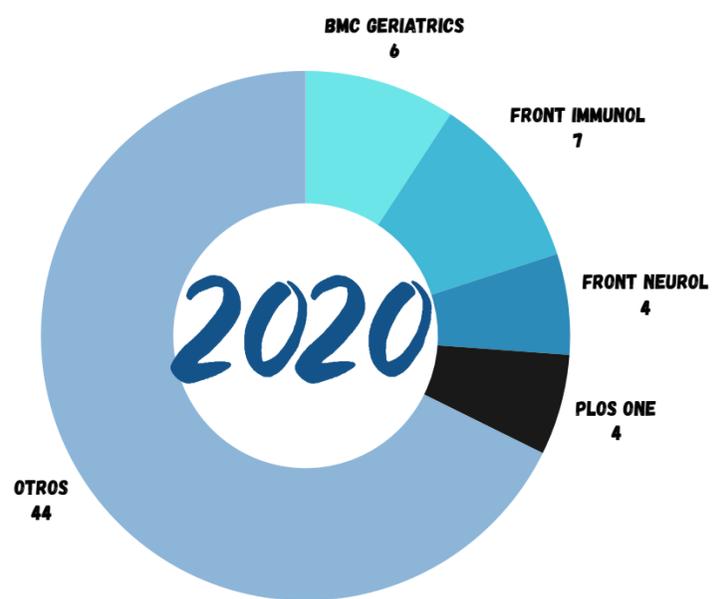


TOTALES TRABAJOS publicados 5 años

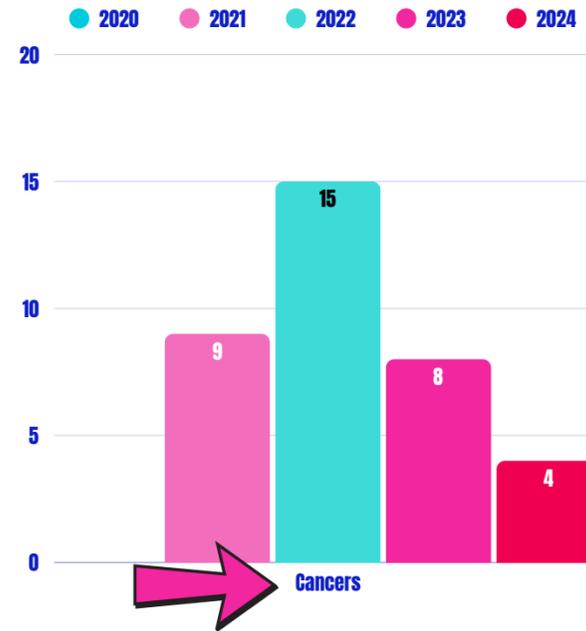
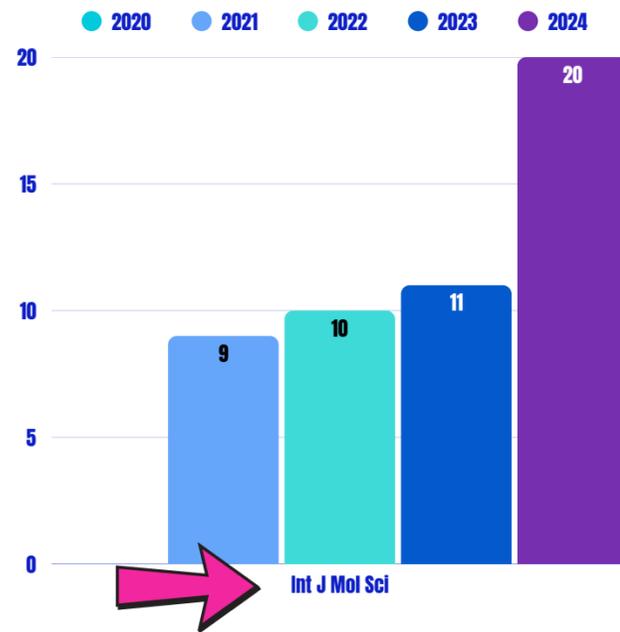


RANKING	REVISTA	ISSN	LICENCIA	EDITOR	DECLARACIÓN DE DATOS
130	J Clin Med	2077-0383	CC BY	MDPI	Sí, 111 (85,4%)
50	Int J Mol Sci	1422-0067	CC BY	MDPI	Sí, 42 (84%)
35	Cancers	2072-6694	CC BY	MDPI	Sí, 31 (88,6%)
35	Nutrients	2072-6643	CC BY	MDPI	Sí, 32 (91,4%)
35	Front Immunol	1664-3224	CC BY	Frontiers Media	Sí, 30 (63,8%)
31	Sci Rep	2045-2322	CC BY	Nature Portfolio	Sí, 30 (96,7%)
21	Biomedicines	2227-9059	CC BY	MDPI	Sí, 21 (100%)
16	Plos One	1932-6203	CC BY	Public Library Science	Sí, 16 (100%)
15	Front Immunol	1664-3224	CC BY	Frontiers Media	Sí, 15 (100%)
13	J Pers Med	2075-4426	CC BY	MDPI	Sí, 13 (100%)
11	Front Oncol	2234-943X	CC BY	Frontiers Media	Sí, 10 (90,9%)
10	Front Neurol	1664-2295	CC BY	Frontiers Media	Sí, 10 (100%)
9	Cells	2073-4409	CC BY	MDPI	Sí, 6 (66,6%)
9	Front Cardiovasc	2297-055X	CC BY	Frontiers Media	Sí, 9 (100%)
9	Int J Environ Res Public	1660-4601 1661-7827	CC BY	MDPI	Sí, 9 (100%)

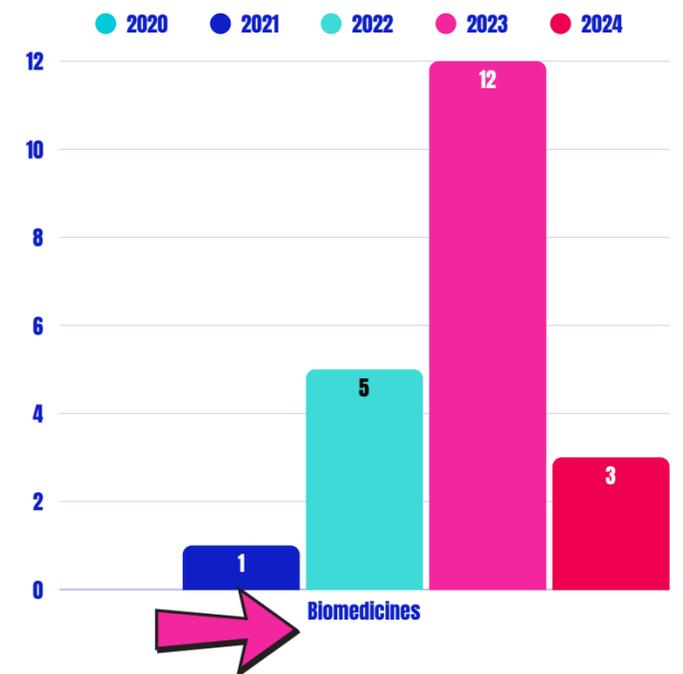
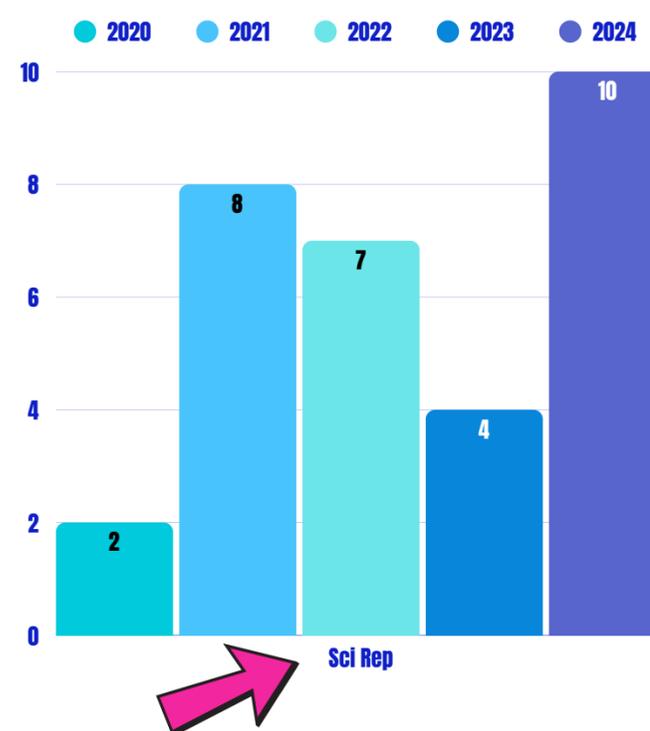
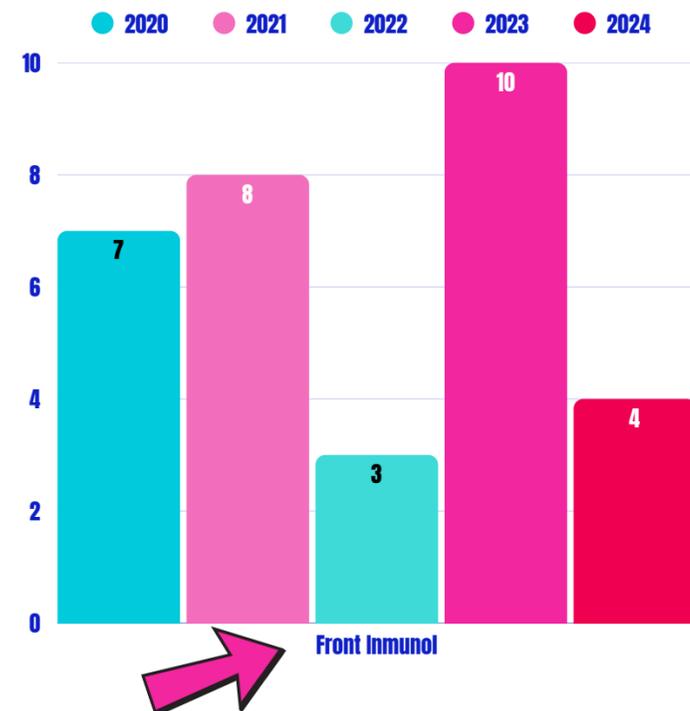
MÉTODO Y RESULTADOS: Análisis cuantitativo resultados - RANKING REVISTAS - Totales años y título revistas (+4) por año



MÉTODO Y RESULTADOS: Análisis cuantitativo resultados - RANKING REVISTAS -



**Título de las Revistas Publicados
+ 20 veces
distribuido a lo largo de los 5 años
(2020-2024)**



MÉTODO Y RESULTADOS: ESTUDIO DE CAMPO - Revisión manual texto completo trabajos -

TIPOLOGÍA MATERIAL ASOCIADO, DECLARACIÓN DE DATOS ¿Disponibilidad? ¿Cómo? ¿Dónde?

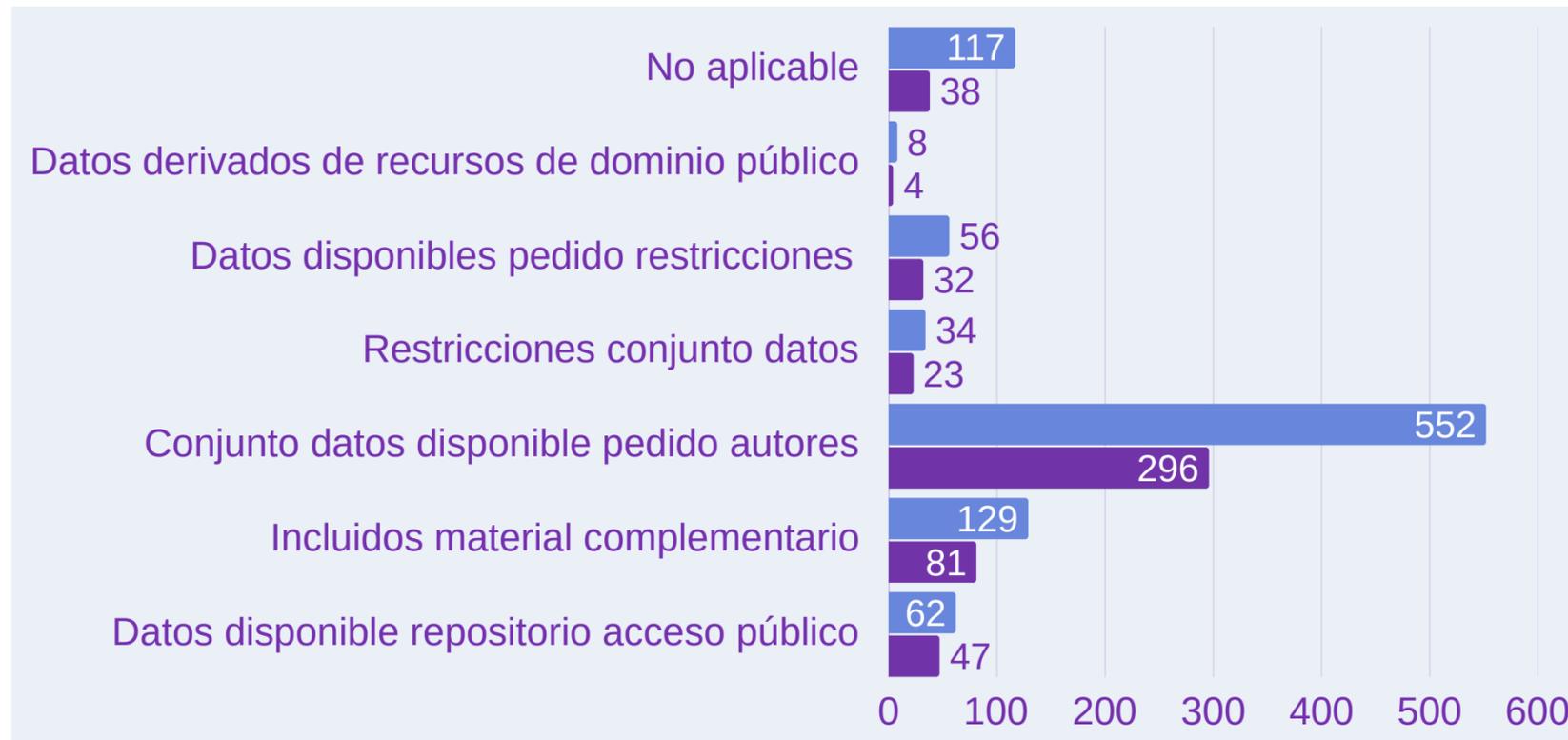


Estado de disponibilidad de los datos	Declaración de disponibilidad de datos recomendada
Datos disponibles en un repositorio de acceso público	Los datos originales presentados en el estudio están disponibles abiertamente en [nombre del repositorio, por ejemplo, FigShare] en [DOI/URL] o [número de referencia/acceso].
Datos disponibles a pedido debido a restricciones (por ejemplo, privacidad, razones legales o éticas)	Los datos presentados en este estudio están disponibles a solicitud del autor correspondiente debido a (especifique el motivo de la restricción).
Datos de terceros	La disponibilidad de estos datos está sujeta a restricciones. Los datos se obtuvieron de [tercero] y están disponibles [de los autores/en la URL] con la autorización de [tercero].
Embargo de datos debido a restricciones comerciales	Los datos que respaldan los hallazgos estarán disponibles en [nombre del repositorio] en [enlace URL / DOI] luego de un embargo a partir de la fecha de publicación para permitir la comercialización de los resultados de la investigación.
Se aplican restricciones a los conjuntos de datos.	Los conjuntos de datos presentados en este artículo no están disponibles fácilmente porque [incluya el motivo, p. ej., los datos forman parte de un estudio en curso o debido a limitaciones técnicas o de tiempo]. Las solicitudes de acceso a los conjuntos de datos deben dirigirse a [entrada de texto].
Datos derivados de recursos de dominio público	Los datos presentados en este estudio están disponibles en [nombre del repositorio] en [URL/DOI], número de referencia [número de referencia]. Estos datos se obtuvieron de los siguientes recursos disponibles en el dominio público: [lista de recursos y URL].
El intercambio de datos no es aplicable (solo es apropiado si no se generan datos nuevos o el artículo describe una investigación completamente teórica)	En este estudio no se crearon ni analizaron nuevos datos. El intercambio de datos no es aplicable a este artículo.
Los datos están contenidos en el artículo o material complementario.	Las contribuciones originales presentadas en este estudio se incluyen en el artículo/material suplementario. Para más información, contacte con el/los autor(es) correspondiente(s).
Conjunto de datos disponible a solicitud de los autores	Los datos brutos que respaldan las conclusiones de este artículo estarán disponibles a solicitud de los autores.

TIPO DE DECLARACIÓN DE DATOS	Volumen/Año	+ Datos Suplementarios	
Conjunto de datos disponible a pedido de los autores	552	296	
	2020		32
	2021		104
	2022		122
	2023		158
2024	136		
Datos disponibles en un repositorio de acceso público	62	47	
	2020		7
	2021		15
	2022		20
	2023		9
2024	11		
Datos derivados de recursos de dominio público	8	4	
	2020		1
	2021		4
	2022		1
	2023		1
2024	1		
Datos disponibles a pedido debido a restricciones (por ejemplo, privacidad, razones legales o éticas)	56	32	
	2020		7
	2021		9
	2022		15
	2023		12
2024	13		
El intercambio de datos no es aplicable	117	38	
	2020		3
	2021		17
	2022		33
	2023		39
2024	25		
Los datos están contenidos en el artículo o material complementario	129	81	
	2020		10
	2021		20
	2022		30
	2023		35
2024	34		
Se aplican restricciones a los conjuntos de datos	34	23	
	2020		3
	2021		8
	2022		13
	2023		8
2024	2		
	958	521	

Tabla 1. MDPI traducción al castellano, de sus recomendaciones para autores sobre cómo declarar los datos de los manuscritos enviados para su revisión.

MÉTODO Y RESULTADOS: ESTUDIO DE CAMPO - **Revisión manual texto completo trabajos -** **RESULTADOS 958 trabajos con Declaración de Datos**



- **Datos no aplicable - diferente tipología artículo -**
- **Datos derivados de recursos dominio público- (INE ,CMBD - Ministerio Sanidad)**
- **Disponible pedido con restricciones (privacidad, ética, etc.)**
- **Se aplican restricciones conjunto de datos- cumplimiento políticas laboratorio, fundación...**
- **PEDIDO AUTORES ?? - sin información datos formatos, plan de gestión, etc.**
- **Contenidos artículo como material suplementario (gran volumen enlace roto)**
- **DATOS REPOSITORIO (base de datos, portales especializados, etc.)**

MÉTODO Y RESULTADOS: ESTUDIO DE CAMPO - Revisión manual texto completo trabajos -

Característica de los trabajos publicados con Declaración de datos

Nº	Denominación	Tipo de plataforma	Acceso/enlace	Cita
1	16S rRNA Database	Base de datos específica para secuencias ARNr 16S de National Center for Biotechnology Information (NCBI).	1, si registro	no
3	BioProject	Colección organizada de datos biológicos relacionados con un proyecto de investigación específico, gestionada por NCBI.	1, si registro	no
3	CBioportal	Plataforma de código abierto diseñada para explorar, visualizar y analizar datos genómicos del cáncer, funciona parte como un repositorio	no	no
6	DIGITAL CSIC	Repositorio institucional multidisciplinar CSIC.	5 con URI 2 con DOI y 1 después publicación	no
1	DRYAD	Repositorio digital de datos de investigación, especializado en biomedicina (trabajo recuperado cita en PMC)	Si, URI, DOI	1, si
5	EMBL-EBI y EGA (European Genome-Phenome)	European Bioinformatics Institute. Laboratorio Europeo de Biología Molecular. Plataforma de Big Data base de datos de secuencias de nucleótidos del EMBL.	3, si	no
2	ENA	European Nucleotide Archive, es una colección global de recursos de biodatos.	2, si	no
1	European Genome-phenome Archive (EGA)	Repositorio europeo seguro que almacena y distribuye datos genómicos y fenotípicos humanos generados por estudios biomédicos.	1, si	no
1	EVA	European Variation Archive base de datos de acceso abierto con todo tipo de datos de variación genética de todas las especies.	1, si	no
2	Figshare	Repositorio digital para datos, artículos, presentaciones, videos, figuras, y otros tipos de contenido académico. Asigna DOI.	2, si DOI	no
1	Fundación ACE	Fundació ACE es una entidad sin ánimo de lucro al servicio de las personas con enfermedad de Alzheimer u otras demencias.	1, no	no
5	GenBank	Base de datos pública de secuencias de ácidos nucleicos (ADN y ARN) mantenida por el NCBI.	5, si	no
9	Gene Expression Omnibus (GEO)	Base de datos que contiene datos de expresión génica y otros datos funcionales, también mantenida por el NCBI. ¹	9, si	no
1	GEPIA2 y Timer.2	(Gene Expression Profiling Interactive Analysis v2) Herramienta webs centrada en el análisis del microambiente inmune tumoral.	1, si	no
1	GitHub	Plataforma alojar proyectos de software usando Git, un sistema de control de versiones ¹¹ .	1, si	no

Nº	Denominación	Tipo de plataforma	Acceso/enlace	Cita
1	Hope Project	Organización global que ofrece atención médica y ayuda humanitaria a comunidades en todo el mundo, con especial énfasis en situaciones de emergencia y crisis.	1, no	no
1	LOVD v.3.0— Leiden Open Variation Database	Base de datos que facilita la recopilación y visualización en línea de variantes de ADN centradas en los genes	1, si	no
1	National Center for Research Resources (NCRR)	Institutos Nacionales de Salud (NIH) que se centró en proporcionar recursos e instalaciones de investigación compartidos para apoyar la investigación biomédica Se disolvió en 2011 ¹¹	1, no	no
1	OA-Breast Database	Conjunto de datos desarrollado por el Computational Imaging Science Laboratory de la Universidad de Illinois ¹² ,	no	no
2	OSF (Open Science Framework)	Repositorio digital por el Center for Open Science (COS), datos sin procesar accesible mediante acuerdo.	si	no
1	PRIDE	Consortio ProteomeXchange a través del repositorio de socios, no en acceso abierto.	1, si	no
1	RIBRON	Registro Español Informatizado de Pacientes con Bronquiectasias	no	no
1	RIMAsalud	Repositorio Institucional de la Consejería de Sanidad de la Comunidad de Madrid	Si, URI	no
5	SCOPE Project	SCOPE tiene como objetivo fomentar la innovación y el crecimiento empresarial en Europa mediante el desarrollo de habilidades de intraemprendimiento	5, no	no
5	Zenodo	Repositorio de acceso abierto multidisciplinar, desarrollado bajo el programa OpenAIRE de la Unión Europea y gestionado por el CERN.	5, URI	1, si
1	No se encuentran en el material suplementario	Los conjuntos de datos presentados en este estudio se encuentran en repositorios en línea. Los nombres de los repositorios y sus números de acceso se encuentran en el artículo/ Material complementario.	1, no	no



Patrones significativos en la gestión de declaración de datos por parte Autores, Editoriales y PMC

DATOS DISPONIBLES, DECLARACIÓN DE DATOS Y BASE DE DATOS

New Try this search in PMC Beta Search

- **Dificultad para localizar de forma automática artículos con DATOS ASOCIADOS y DECLARACIÓN de DATOS**, más adecuado FILTRO PMC, he incluso al revisar manualmente los trabajos localizamos cambios en esta información.
- **Declaraciones de datos, NO VERIFICADO ??** completadas de forma técnica y conforme a los requerimientos pero sin que tengan un concepto de que significa o que información deberían aportar.

PUBLICACIONES Y DISPONIBILIDAD DE DATOS

- **47,5%** Resultados totales PMC tienen **datos asociados**. Refleja esfuerzo autores por cumplir políticas OA, últimos 5 años.
- **52,5% No** (gran volumen), eliminar tipología trabajo, siguen **barreras éticas, limitaciones técnicas o una falta de infraestructura adecuada para la planificación, gestión y almacenamiento de grandes conjuntos de datos.**
- De los **962** con Declaración de datos, el **57,6% los Datos disponibles previa solicitud a los autores**, aunque los autores están dispuestos a compartir sus datos, persisten restricciones significativas en **cuanto al acceso abierto total.**

MATERIAL SUPLEMENTARIO Y CITACIÓN DE DATOS

- **48,6% Artículos con datos disponibles incluyen material suplementario**, uso autores para proporcionar información adicional
- **13,5%** de los trabajos con datos asociados declara incluirlos en el **Material suplementario** dentro del artículo
- **Solo 2** de los artículos analizados incluyera una citación formal de los datos (uno de ellos incluso no rescatado en la búsqueda bibliográfica), omisión crítica en términos de buenas prácticas científicas, especialmente en la inclusión de los datos en un Repositorio (facilita trazabilidad y reutilización, limita visibilidad e impacto).

DISCUSIÓN II



Patrones significativos en los patrones de declaración de datos por parte Autores, editoriales y PMC

IMPACTO DE LAS REVISTAS Y POLÍTICAS EDITORIALES

- **MDPI** editorial con más trabajos publicados, facilita en su política editorial una lista de estados de disponibilidad de datos, mejorares prácticas de ciencia abierta y la transparencia de su investigación y difusión. Permite seguir las *Transparency and Openness Promotion Guidelines (TOP)*, muy especialmente la transparencia de los datos.
- **Todas publicaciones vía dorada de acceso abierto, creadas con el movimiento de OA**, mejorar Declaración, no solo estandarizado si no corroborando su cumplimiento de buenas prácticas y el uso de identificadores persistentes tipo URI, DOI, etc. que faciliten recuperación del material que debe aportar para su verificación

LIMITACIONES DE ESTUDIO

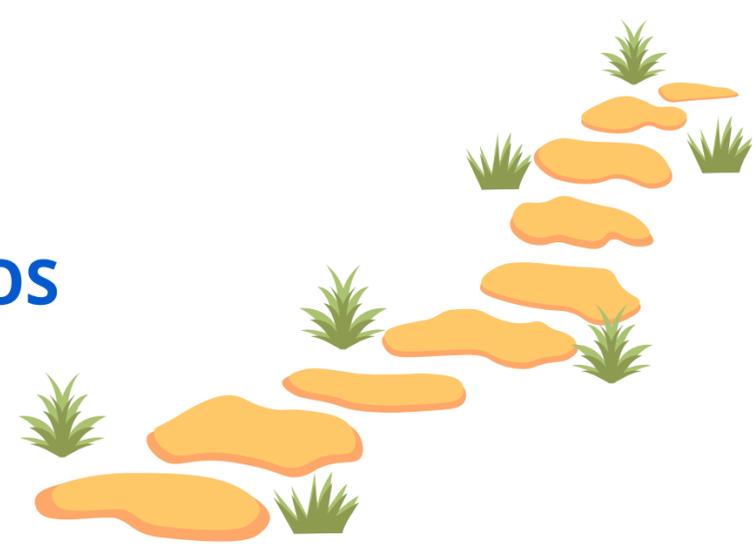
- **Solo** un centro hospitalario.
- **PMC contenido sujeto a acuerdos editoriales** para incluir revistas fomentar OA y preservación. Ej. MDPI incorpora 92 revistas a PMC de las 471 que publica y Frontiers Media 60 de las 230 que publica.

CONCLUSIÓN

A pesar esfuerzo cumplimiento políticas **OA, AUTORES, EDITORES Y BASES DE DATOS**

Largo Camino... con áreas de mejora:

- **Estandarización y verificación** de políticas editoriales sobre OA (Declaración datos, citación, uso identificadores persistentes, gestión datos, uso Repositorios).



BIBLIOGRAFÍA

- Data Filters in PMC and PubMed. NLM Tech Bull. 2018 Mar-Apr;(421):b8. Disponible en: https://www.nlm.nih.gov/pubs/techbull/ma18/brief/ma18_pmc_data_filters.html .
- BOE. 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. Disponible en: <https://www.boe.es/eli/es/l/2022/09/05/17/con>
- Estrategia Nacional de Ciencia Abierta 2023-2027. Disponible en: <https://www.ciencia.gob.es/Estrategias-y-Planes/Estrategias/ENCA.html>
- Ley Orgánica 2/2023, del Sistema Universitario Disponible en: <https://www.boe.es/eli/es/lo/2023/03/22/2/con>
- Plan Estatal de Investigación Científica y Técnica y de Innovación (PEICTI).Disponible en: <https://www.ciencia.gob.es/Estrategias-y-Planes/Planes-y-programas/PEICTI.html>
- IdISSC ¿Cómo citar correctamente en las publicaciones? Disponible en: <https://www.idissc.org/wp-content/uploads/2022/01/2019-03-25-Como-citar-las-publicaciones.pdf>).
- Recomendaciones Política Editorial MDPI a los autores sobre como realizar la “Declaración de disponibilidad de los datos”. Disponible en: <https://www.mdpi.com/journal/ijerph/instructions#suppmaterials>
- Quaderi N. Supporting integrity of the scholarly record: Our commitment to curation and selectivity in the Web of Science Disponible en: <https://clarivate.com/academia-government/blog/supporting-integrity-of-the-scholarly-record-our-commitment-to-curation-and-selectivity-in-the-web-of-science/>
- Recomendaciones Política Editorial MDPI sobre “Estado o disponibilidad de los datos” de la revistaInt J Environ Res Public Health. Disponible en: <https://www.mdpi.com/journal/ijerph/instructions#suppmaterials> .
- New Measure Rates Quality of Research Journals’ Policies to Promote Transparency and Reproducibility. Center for Open Science,. Disponible en: <http://www.cos.io/about/news/new-measure-rates-quality-research-journals-policies-promote-transparency-and-reproducibility> (acceso 04/08/2024)
- MDPI journal list. Disponible en: <https://www.mdpi.com/about/journals>
- PMC. Selective Deposit Colection. Disponible en: <https://pmc.ncbi.nlm.nih.gov/about/selectivedeposit/>



*Muchas
a Gracias*

**Biblioteca de Ciencias de la Salud
Hospital Clínico San Carlos**



Blanca San José Montano
blanca.sanjose@salud.madrid.org