

Material suplementario: Datos y código de reproducibilidad

Héctor García-Leal ¹ y Álvaro Puelles-Díaz ²

¹Universidad de La Serena, Chile

²Universidad de La Serena; Universidad del Alba, Carrera de Medicina, Chile

Costo de publicar en ciencias de la salud en Chile: análisis de cargos por procesamiento de artículos mediante datos abiertos de OpenAlex, 2019–2024

Revista Internacional de Ciencia Abierta

<https://revistas.userena.cl/index.php/cienciaabierta>

Volumen 2 **Artículo:** <https://doi.org/10.15443/rintca.2026.art001>

Depósito de datos (Zenodo): <https://doi.org/10.5281/zenodo.20790843>

OPEN ACCESS — CC BY 4.0

1. Descripción del depósito

Este documento da cuenta del material suplementario asociado al artículo, depositado de forma abierta y permanente en Zenodo. El depósito contiene el conjunto de datos completo, el código en R que lo genera y procesa, y los productos de análisis (tablas y figuras) que sustentan los resultados publicados. Su objetivo es garantizar la reproducibilidad independiente del estudio conforme a las prácticas de ciencia abierta.

Identificación del depósito

Repositorio	Zenodo (CERN / OpenAIRE)
DOI (versión)	https://doi.org/10.5281/zenodo.20790843
Tipo	Conjunto de datos (<i>Dataset</i>)
Licencia	Creative Commons Attribution 4.0 International (CC BY 4.0)
Relación	<i>Is supplement to</i> → artículo DOI 10.15443/rintca.2026.art001

2. Contenido del depósito

El depósito reúne catorce archivos organizados en código, datos, tablas y figuras.

Archivo	Descripción
<i>Código R</i> 01_extraccion_APC_salud_2019_2024.R	Extracción desde OpenAlex y análisis; genera tablas y dataset.
02_figuras_APC_salud.R	Genera las tres figuras del artículo.
<i>Datos</i> dataset_APC_salud_chile_2019_2024_snapshot_2026-06-05.csv obras_crudo_2019_2024.rds	Dataset limpio (23.230 publicaciones, 10 variables), en formato abierto CSV. Snapshot crudo congelado de OpenAlex (ancla de reproducibilidad).
<i>Tablas de resultados (.xlsx)</i> tabla1_indicadores_globales tabla2_distribucion_anual tabla2_con_sensibilidad tabla3_editoriales tabla4_afiliaciones tabla_revistas_top15	Indicadores globales del gasto APC. Distribución anual del gasto. Distribución anual con análisis de sensibilidad por cobertura. Concentración del gasto por editorial (HHI, CR3). Afilaciones institucionales más frecuentes. Revistas con mayor gasto APC acumulado.
<i>Figuras (.png)</i> figura1_evolucion figura2_editoriales figura3_afiliaciones	Evolución del gasto y tasa de registro. Concentración del gasto por editorial. Afilaciones institucionales.
README.md	Documentación: metodología, estructura y orden de ejecución.

3. Resumen metodológico

Parámetros de extracción (snapshot reproducible)

Fuente	OpenAlex API (post-Walden)
Paquete / lenguaje	openalexR v2.0.0 / R v4.5.2
Clasificación disciplinar	Topics — Field "Medicine" (id 27)
Filtro geográfico	authorships.institutions.country_code = "CL"
Tipo de obra y período	article; 2019-01-01 a 2024-12-31
Campo de gasto	apc (subcampo type="paid", value_usd)
Fecha de snapshot	5 de junio de 2026 (2026-06-05)
Deduplicación	DOI único (global); par DOI×institución (afiliaciones)

Cifras de referencia. 23.230 publicaciones recuperadas; 5.944 (25,6%) con APC registrado como pagado; gasto total estimado USD 13.988.171 (promedio USD 2.353; mediana USD 2.382; máximo USD 11.690).

4. Reproducibilidad

Para reproducir las cifras exactas del estudio se recomienda **no reextraer** desde la API, sino cargar el snapshot crudo archivado, dado que OpenAlex se actualiza de forma continua y una nueva consulta puede arrojar resultados distintos:

```
obras <- readRDS("datos_crudos/obras_crudo_2019_2024.rds")
```

A partir de ese objeto, el script 01_extraccion_APC_salud_2019_2024.R regenera las tablas y el dataset, y 02_figuras_APC_salud.R regenera las figuras. El archivo README.md detalla el orden de ejecución.

5. Cómo citar el depósito

García-Leal, H., & Puellas-Díaz, Á. (2026). *Datos y código de reproducibilidad: Costo de publicar en ciencias de la salud en Chile (APC, OpenAlex 2019–2024)* [Conjunto de datos]. Zenodo. <https://doi.org/10.5281/zenodo.20790843>

Este documento describe el material suplementario del artículo <https://doi.org/10.15443/rintca.2026.art001>. Los datos y el código se distribuyen bajo licencia CC BY 4.0.