

¡Cuartiles! ¡Ay mis cuartiiiileees! (Canción popular)

Quartiles! Oh, my quartiiiileees! (Popular song)

Antonio Calderón-Rehecho

Calderón-Rehecho, Antonio (2021). "¡Cuartiles! ¡Ay mis cuartiiiileees! (Canción popular)". *Anuario ThinkEPI*, v. 15, e15e02.

<https://doi.org/10.3145/thinkepi.2021.e15e02>

Publicado en *IweTel* el 17 de mayo de 2021

Antonio Calderón-Rehecho

<https://orcid.org/0000-0002-9948-2825>

Universidad Complutense de Madrid
Biblioteca
acaldero@ucm.es



Resumen: Se analizan los aspectos implicados en el cálculo del factor de impacto y su aplicación en la elaboración de rankings, tomando ejemplos del IDR de *Dialnet Métricas*. Se pone de manifiesto su sinsentido para valorar autores y hasta las mismas revistas, recomendando otras perspectivas, como las que aporta *Indicadores Dialnet*.

Palabras clave: Factor de impacto; Rankings de revistas; Evaluación de la investigación; *Dialnet Métricas*.

Abstract: The problematic associated to impact factor (and the rankings create with it) is analyzed by examples in the *Dialnet Métricas'* IDR. Its nonsense for evaluating researchers or journal is revealed and it is stated that other criteria should be considered, as those included in *Dialnet* indicators.

Keywords: Impact factor; Journal rankings; Research evaluation; *Dialnet Métricas*.

1. Introducción

No era mi idea original para un nuevo *ThinkEPI*, pero se vio trastocada por el extraño 2020. Ni era mi intención tratar este tema ya que doctores tiene la Iglesia... (por ejemplo, **Delgado-López-Cózar; Martín-Martín**, 2019; **Aguillo**, 2021), pero algunas reacciones ante la nueva edición de *Dialnet Métricas* (2021) y la noticia (*Ministerio de Ciencia e Innovación*, 2021) de que la *Agencia Estatal de Investigación* parece haberse adherido a *DORA* (**Codina**, 2019), me han animado a hacerlo.

Independiente de si su sola mención nos asusta, nos molesta o nos obsesiona, el cometido de los cuartiles es indicarnos una posición relativa dentro de un listado ordenado, sin entrar en consideraciones sobre qué es lo que se ordena.

Podría utilizarse en listados que resultan de cálculos complejos que utilizan muchas variables, ponderaciones rigurosas... Imaginemos una lista de países ordenados según el IPC (en el hipotético caso de que se hubiera conseguido que más de uno comparta los mismos criterios para establecer el índice) o dentro de nuestro ámbito disciplinar, que *MIAR* estableciera un listado ordenado según la puntuación (ICDS) que otorga tras considerar varios parámetros, o que hiciéramos lo propio con la clasificación *CIRC* (algo que ya realiza en cierto modo ella misma).

Sin embargo, vamos a hacerlo más sencillo, centrándonos en los que se utilizan para establecer la posición en un listado ordenado según el factor de impacto de las revistas. Y ya que nos ha servido de inspiración, partiremos de ejemplos y consideraciones del *Índice Dialnet de Revistas*.

2. Cuartiles

Los cuartiles dividen un listado ordenado en 4 partes iguales, de tal manera que, si tuviéramos 100 unidades, cada uno de los cuartiles contendría 25 unidades (1/4 del total). Como ese listado está ordenado, el último elemento del primer cuartil tendrá un valor más alto que el primer elemento del segundo y así sucesivamente.

Hay que matizar esta afirmación, dado que en ocasiones el número que se obtiene al dividir entre 4 no es entero. Por ejemplo, si en el listado hubiera 38 elementos, el cuartil debería dividirse en el elemento 9,5. Ante esta situación, se pueden tomar diferentes decisiones; por ejemplo, hacer que tengan 19 los dos primeros cuartiles y los dos últimos, repartidos: 9, 10, 9, 10. O incluir directamente los 2 que causan la discordia al final, con el resultado de 9, 9, 9, 11, o...

También puede suceder que los últimos integrantes de un cuartil ideal y los primeros del siguiente tengan los mismos valores.

O que, para establecer el reparto, no tengamos en cuenta las revistas que no tienen citas y las acumulemos al final. Si tenemos 40 revistas, de las que 4 no tienen citas, se considerarían 36 para obtener el cuartil. En tal caso los tres primeros cuartiles tendrían 9 revistas cada uno y el último, 13. La figura 1 muestra el resultado en un listado de 16 revistas con 4 sin citas.

La misma idea se aplica a deciles y percentiles, solo que en su caso las subdivisiones son diferentes: el listado se divide en 10 y en 100 partes iguales respectivamente.

Resulta curioso que tanto cuartiles como deciles siguen el criterio ordinal para indicar el que ocupa una mejor posición, pero en el caso de los percentiles, el que se considera mejor situado tiene el número mayor: primer cuartil (Q1, C1), segundo cuartil (Q2, C2), primer decil (D1), segundo decil (D2)... pero P99, P98, P97...

3. Factor de impacto: elementos

El factor de impacto de las revistas es un índice obtenido de dos valores: el número de citas recibidas y el número de documentos publicados; ambos en un período de tiempo determinado. El clásico de Garfield usaba 2 años, el índice de impacto del IDR utiliza 5. Por pura lógica el impacto obtenido será mayor cuantas más citas se obtengan y se verá beneficiado por un número menor de artículos publicados.

En la figura 2 hemos ordenado los datos de las revistas más citadas de un ámbito temático por orden descendente de citas.

Si nos fijamos en los datos de impacto, las dos primeras están en el orden que les corresponde. El resto no; de hecho, la que aparece en último lugar es la que tiene el tercer impacto más alto.

Las dos primeras tienen un índice de impacto mucho más relevante que el resto, sin embargo, tampoco hay tanta diferencia entre ellas a pesar de que la primera tiene casi 4 veces más citas (pero también tres veces más artículos).

Si nos centramos en las dos últimas, la diferencia de citas es mínima (4), sin embargo, el índice de impacto de la última, la menos citada, es el doble. De hecho, la penúltima está en un cuartil diferente.

Y por mencionar otra cuestión más que apunta a la concentración de citas en pocas revistas, la primera multiplica por más de 10 el número de artículos de la última, pero consigue casi 29 veces más citas. De hecho, tiene prácticamente el doble de citas que la suma de todas las demás. Este aspecto supone que pequeñas variaciones en el número de citas conlleven el cambio de posición en el ranking.

En rigor realizar este cálculo y ordenar según los resultados no tiene complicación, ya que la esencia es contar citas y contar documentos. Pero la realidad es diferente, como veremos enseguida.

Crtl	Pctl	# ↑
C1	P99	1
C1	P84	2
C1	P76	3
C2	P69	4
C2	P61	5
C2	P53	6
C3	P46	7
C3	P38	8
C3	P30	9
C4	P23	10
C4	P15	11
C4	P07	12
C4	P00	13
C4	P00	13
C4	P00	13
C4	P00	13

1 - 16 de 16

Figura 1. Listado de 16 revistas, 4 de ellas sin citas.

≡ Impacto 5 años	≡ Artículos	↓ Citas
1,448	493	714
1,132	167	189
0,310	216	67
0,375	128	48
0,257	113	29
0,532	47	25

Figura 2. Revistas ordenadas por número de citas

4. Sobre los universos: revistas, artículos y citas

4.1. El universo de revistas

De todos es conocido que no todas las revistas están en *Web of Science (WoS)*. Es un objetivo a conseguir por casi cualquiera de ellas. Pero, además, no todas las revistas que están en *WoS* aparecen en los *JCR*. Ni siquiera hay *JCR* para Arte y Humanidades.

Significa esto que para estar en el listado sobre el que se determinan los cuartiles se precisa una selección que acote el universo sobre el que se trabaja. La selección puede ser más o menos rigurosa, pero siempre existe y es modulable: el comportamiento dentro del esquema establecido puede hacer que algunas se queden fuera en ediciones posteriores.

Por ejemplo, en el IDR la selección tiene en cuenta a revistas españolas científicas o técnico-profesionales, que sean de Ciencias Sociales, Humanidades o Derecho, que encontraremos en *Latindex*, en *WoS*, en *Scopus*, en los listados de *Google Scholar Metrics*, en los antiguos *Inrexx* o en los listados de la *Fecyt*.

Una segunda selección tiene que ver con el ámbito temático al que pertenecen. Puesto que los comportamientos son diferentes dependiendo de las disciplinas, es preciso acotar adecuadamente las revistas que aparecerán en un ranking u otro. Imaginemos el efecto que tendrían unas pocas revistas de Química Analítica en rankings de Arte o Filología.

Claro está que, por mucho que cada uno de nosotros consideremos que los límites de nuestro ámbito de estudio están suficientemente claros y asentados, no hay compartimentos estancos ni todo curriculum que se nos ocurra podrá ser aislado de los demás.

Además, y recojo esta idea de Emilio Delgado en una jornada de la UNIR en diciembre de 2019 (**Delgado-López-Cózar**, 2019), las revistas no son monotemáticas.

Un ejemplo de la importancia de este universo lo podemos observar con relativa facilidad con alguno de los tres ámbitos genéricos de *Dialnet Métricas*: Historia, Filología y Derecho. Cada uno de ellos engloba las revistas de subdivisiones menores de la disciplina que tienen sus rankings específicos. Hemos escogido Derecho, ya que ha sido el que más ha tardado en llegar.

Tabla 1. Subdivisiones de Derecho

Ámbito	C1	C2	C3	C4
Derecho Administrativo	33,3%	27,8%	13,9%	25,0%
Derecho Civil y Mercantil	11,8%	23,5%	27,5%	37,3%
Derecho Constitucional y Político	45,0%	25,0%	15,0%	15,0%
Derecho del Trabajo	32,4%	29,4%	14,7%	23,5%
Derecho Eclesiástico	30,0%	10,0%	40,0%	20,0%
Derecho Financiero y Tributario	25,0%	25,0%	20,0%	30,0%
Derecho Internacional	44,4%	14,8%	22,2%	18,5%
Derecho Multidisciplinar	11,5%	23,0%	24,1%	41,4%
Derecho Penal, Procesal y Criminología	34,2%	21,1%	28,9%	15,8%
Derecho Romano e Historia del Derecho	0,0%	23,8%	28,6%	47,6%
Filosofía del Derecho	27,8%	27,8%	27,8%	16,7%

En la tabla 1 se observan 11 subdivisiones de Derecho, seguramente discutibles, posiblemente insuficientes, dependerá del punto de vista de cada cual. Básicamente son las mismas que estaban en *Inrexx*.

Cada una de ellas tiene su IDR particular con universos más pequeños y especializados. Pero comparten un ranking en el que se incluyen todas. Lo que hemos hecho ha sido comprobar qué porcentaje de las revistas de cada subdivisión se encuentra en cada uno de los cuartiles del ranking general. Y hemos marcado en azul claro los datos que coinciden con el 25% (en un reparto equilibrado entre cuartiles sería lo que correspondería a cada uno) y con azul oscuro los que superan esa cifra.

Los datos hablan por sí mismos. Apuntamos alguno:

- Derecho Romano e Historia del Derecho no tienen ninguna revista en el primer cuartil. Sin embargo, en su ranking particular indudablemente la tienen.
- Civil y Mercantil se unen a él y a Multidisciplinar concentrándose en los últimos cuartiles.
- Romano e Historia y Multidisciplinar tienen más del 40% de las revistas en el último cuartil. Multidisciplinar es el que más revistas incluye.
- En el sentido contrario, superan el 40% en el primer cuartil Internacional y Constitucional y Político.

- Administrativo y Trabajo los acompañan superando el 60% en los dos primeros cuartiles.
- Financiero y Tributario casi consigue un reparto equilibrado entre cuartiles.

Corolario: las posiciones cambian dependiendo de con quién nos comparemos. Por eso es importante que lo hagamos con los iguales.

Si un ámbito tiende a concentrar más citas que otro, sus revistas tenderán a ocupar una posición superior cuando se unan en un ranking.

Corolario azul-oscuro-casi-negro: cuantas menos revistas haya, más posibilidades de estar mejor situado. O menos, porque también hay menos puestos en cada cuartil y en el supuesto absurdo de que solo hubiera una revista, esta ocuparía los 4 cuartiles.

“Las posiciones cambian dependiendo de con quién nos comparemos. Por eso es importante que lo hagamos con los iguales”

4.2. Los artículos citables

Como dice **Delgado-López-Cózar** (2019), las revistas son conglomerados de documentos.

No todos los documentos individuales de una revista deberían ser considerados de la misma manera a la hora de recibir citas: las cartas al director, las reseñas o similares (de cualquier tipo: las hay de tesis doctorales, de discos, de jurisprudencia, etc.), las entrevistas, las creaciones artísticas, los obituarios, las columnas de opinión, la información sobre productos, la propaganda, las bibliografías, las fichas, las guías, las conversaciones, las crónicas, las noticias, los retratos, los listados de evaluadores, ¿los editoriales?... Establecido el criterio, que debe ser igual para todas, hay que marcar los que deben ser contados y los que no. Y también debería haber una correspondencia en el sentido de que, si un documento no puede ser citado, tampoco pueda citar. Por ejemplo, no tiene sentido considerar (para el cómputo del impacto) las 13 citas de una reseña de apenas dos páginas.

Si como hemos vislumbrado, y veremos a continuación, un número escaso de citas tiene mucha repercusión en el impacto, la reducción del número de documentos citables actúa en el mismo sentido, pudiendo darse el caso de que una revista con escasas citas pero muy pocos artículos publicados aparezca muy arriba en el ranking. Para evitarlo sería conveniente establecer un umbral mínimo de publicación. Es verdad que hacerlo afectaría a las revistas que sacan pocos números al año o con escasos artículos.

Fijándonos en la figura 2, puede parecer que 493 artículos son muchos, pero hay que considerar que estamos hablando de 5 años, lo que supone menos de 100 artículos al año. Si la revista es trimestral, el número queda reducido a 25 artículos por número.

Si hacemos lo mismo con la de 47 artículos, resulta que cada año tendría 9 o 10 artículos, lo que hace suponer que sería un anuario.

También puede ocurrir que la razón de un número escaso sea la no publicación de algún número en alguno de los años de la ventana de citación, o que la revista haya muerto en ese intervalo de tiempo; o, por el contrario, que sea muy nueva. En esos casos en el IDR el dato aparecerá con un asterisco.

4.3. Las citas

Como hemos dicho, siempre hay una selección. Hay sistemas en los que hecha la misma se extraen todas las citas emitidas por las revistas; pero en otros se escoge una muestra (una segunda selección) para utilizar sus citas. Normalmente todas quieren ser parte de esta muestra. Resulta curioso, ya que lo que importa es ser citado y no citante. Seguramente tiene que ver con un aspecto interesante: del mismo modo que los autores se autocitan, también hay una tendencia en las revistas a autocitarse: es lo que se llama autorreferenciación. Hasta cierto punto puede ser lógico, tal y como sucede con las autorías; pero ¿realmente es explicable que el 44% de las citas de una revista provengan de ella misma? Tiene todos los síntomas del dopaje, que “paradójicamente” podría no ser necesario para ocupar puestos importantes en el ranking. Seguramente habría que establecer también un umbral para la tasa de autorreferenciación de tal manera que se excluyera la revista del ranking, o se realizara uno en el que se excluyeran todas las autorreferencias, o...

Si esto se suma a la realidad de que (lo veremos con datos enseguida) muy pocas citas tienen un impacto muy alto, la distorsión es clara.

En el uso de las citas no solo está implicada la selección establecida, sino que a veces son las propias revistas las que se autoexcluyen. Si las revistas solo se publican en papel es complicado que se pueda contar con sus citas. Si no publican las referencias al final, sino solo en notas al pie, igual. No es incompatible incluir notas al pie y referencias al final como lo demuestran un buen número de revistas. Hay que destacar que incluso *Google Scholar* (2021), que cuenta con medios de procesamiento enormes, lo recomienda o considera necesario para poder hacer su trabajo indizador.

En ocasiones también las restricciones de acceso a la bibliografía por la necesidad de suscripción lo complican. Como lo hacen las plataformas que establecen una excesiva “seguridad”.

5. Resultados del impacto

Vamos a utilizar algún ejemplo que nos ayude a entender la realidad (figura 3). En primer lugar, hemos escogido el ranking de una disciplina y hemos incluido el impacto de las revistas que ocupan el primer y último puesto en cada cuartil.

- La revista situada en primer lugar en el ranking tiene un impacto de 0,325 lo que viene a decir que la probabilidad (si se distribuyera uniformemente entre todos los artículos y si mantuviera ese valor en el futuro) de uno de sus artículos de obtener cita sería de casi un tercio: son citas esperadas como dice **Aguillo** (2021, diapositiva 11). O, dicho de otra manera, más del 66% de los artículos de la primera revista del ranking no recibirá ninguna cita.
- Siguiendo el mismo criterio, más del 93,6% de los artículos de la mitad menos citada de las revistas (dos últimos cuartiles) no obtendrán citas.
- Seguramente estamos ante una disciplina para la que las revistas no son la primera fuente de citación, sino que lo será otro tipo de documentos. Incluso podría ocurrir que las propias revistas tuvieran más citas de esos tipos documentales que de las revistas.
- En este listado hay 28 revistas que no han obtenido cita. Casi el 21%. Cabe analizar la idoneidad de incluirlas.
- Podríamos incidir en la escasa diferencia en las fronteras entre cuartiles (una o dos milésimas) pero lo vamos de dejar para el siguiente ejemplo.

Mostramos ahora otra disciplina en la que también escogemos las fronteras entre cuartiles, con el impacto, los documentos y las citas (figura 4). Concretamente el cambio entre el segundo y el tercer cuartil y entre este y el cuarto:

- La variación en el impacto es mínima: 9 milésimas en las situadas más arriba y 2 milésimas en las dos inferiores. ¿Tiene sentido calcular 3 decimales? ¿Tiene sentido decir que una revista está por dos milésimas en un cuartil más bajo que el requerido por un evaluador? Si solo se sacaran dos decimales el resultado sería el mismo.
- De hecho, la variación entre el final del C2 y el comienzo del C4 es de 0,072.
- La revista situada en el último cuartil tiene más citas que la situada en el C2: sólo 2, pero el 20% con respecto al valor absoluto.
- Decir que una revista tiene un impacto de 0,217 sería como decir –si el reparto fuera equitativo entre todos– que de cada 5 artículos uno será citado. O, lo que es lo mismo, que casi el 80% de los artículos no serán citados. En rigor, el 78,3%. En el caso del último dato, el 85,4% de los artículos no se citarían. Como sabemos que las citas se concentran en unos pocos, hace pensar que el porcentaje de no citados es mayor.
- En clave *Dialnet*: los de la primera tabla, en la que se incluyen asteriscos, tienen alguna incidencia en la publicación de los documentos.

Cuartil	Ranking	Impacto
C1	1	0,325
	26	0,121
C2	27	0,119
	53	0,064
C3	54	0,062
	80	0,034
C4	81	0,033
	134	0

Figura 3.

	Impacto 5 años	Documentos	Citas
C2	0,217	46 *	10
C3	0,208	106 *	22
C3	0,148	54	8
C4	0,146	82	12

Figura 4.

6. Siempre nos quedará...

Se habrá observado que no hemos entrado en discusiones teóricas y más bien hemos pretendido mostrar una realidad con ejemplos reales (aunque no hayamos dado nombres) que nos hagan ver que la cuestión de en qué cuartil se situaría una revista no debería ser la que determine el éxito o fracaso de una investigación publicada en ella, sino la propia investigación. Una afirmación que no necesitaría ni de los ejemplos.

Claro que quien no sea el autor de un artículo sino alguien que trabaja para mantener la revista puede pensar de otra manera. Pero ¿realmente las diferencias mínimas que se establecen entre la mayor parte de ellas y los umbrales de impacto tan reducidos que tienen en algunos casos son significativas para afirmar cualquier cosa? O la constatación de que en determinadas disciplinas una o unas pocas revistas concentran el impacto y las citas, ¿no colocan a las demás en situación cuando menos apurada?

“La cuestión de en qué cuartil se situaría una revista no debería ser la que determine el éxito o fracaso de una investigación publicada en ella, sino la propia investigación”

¿No será mucho mejor disponer de toda otra serie de datos que reflejen la realidad de la revista? ¿Incluyendo su devenir histórico?

Por ejemplo:

- afinidad entre revistas según las citas de unas y otras;
- artículos más citados;
- autores más citados (que pertenecen a áreas del conocimiento e instituciones concretas);
- autores citantes (que pertenecen a áreas del conocimiento e instituciones concretas);
- revistas citantes;
- artículos citantes;
- diferentes indicadores como número de citas recibidas, el índice h, el índice h5, la mediana h...

IDR

Trabajos de Prehistoria

0082-5638, 1988-3218



[Ver Indicadores Dialnet](#)

Año	Impacto	Nº citas	ARQUEOLOGÍA Y PREHISTORIA		HISTORIA	
			Posición	Cuartil	Posición	Cuartil
2019	0,899	89	1 / 64	C1	1 / 313	C1

Trabajos de Prehistoria

ISSN: 0082-5638, 1988-3218 País de edición: ESPAÑA Ámbito: HISTORIA, ARQUEOLOGÍA Y PREHISTORIA



C1 IDR 2019 ARQUEOLOGÍA Y PREHISTORIA

C1 IDR 2019 HISTORIA

Indicadores Dialnet

Se calculan a partir de las citas emitidas por todos los artículos que en Dialnet tienen incluidas las referencias bibliográficas

- [Revistas relacionadas](#)
- [Artículos más citados](#)
- [Autores más citados](#)
- [Autores citantes](#)
- [Revistas citantes](#)
- [Artículos citantes](#)

#	Revista	Ámbitos	Afinidad
1	Complutum	ARQUEOLOGÍA Y PREHISTORIA HISTORIA	0,230

Figura 5. Ejemplo de revista en *Indicadores Dialnet*

Todos ya están disponibles, o lo estarán en la edición de 2020, en *Indicadores Dialnet*, la parte de *Dialnet Métricas* más enriquecedora y en la que menos se piensa.

<https://dialnet.unirioja.es/metricas/indicadores>

En *Indicadores Dialnet* además, se pueden ver los documentos más citados en cada disciplina, sean libros, capítulos de libros o tesis. Y, también los investigadores más citados, que se agrupan por áreas de conocimiento y universidades. Cada investigador cuenta con una página específica siempre que tenga más de 4 citas, que proporciona una visión muy completa de sus datos de investigación.

Un espacio no limitado a las citas de las revistas fuente del IDR, donde se experimenta con indicadores y presentaciones gráficas, donde la información se presenta a nivel de autor y de documento y del entorno en que viven tanto uno como otro. Y totalmente interconectado con *Dialnet*. Por eso, para hacernos una idea más real de nuestro perfil, seamos revista, documento o autor siempre nos quedará... Indicadores Dialnet.

7. Referencias

Aguillo, Isidro F. (2021). "Indicios de calidad. Métricas". En: *Jornada sobre la carrera investigadora. Módulo II. Sesión 3: Impacto bibliométrico de la investigación (en línea)*. Facultad de Veterinaria de la UCM, 18 de febrero de 2021, pp. 1-33. <https://veterinaria.ucm.es/file/impacto-bibliometrico-investigacion>

Codina, Lluís (2019). "Manifiesto de Leiden y DORA: situemos las métricas en su lugar". *Lluís Codina*, 31 mayo. <https://www.luiscodina.com/leiden-manifiesto-dora-metricas>

Delgado-López-Cózar, Emilio (2019). "El factor de impacto de una revista: un indicador inadecuado para medir el rendimiento científico que crea incentivos perversos y efectos dañinos para la ciencia". En: *Jornada sobre evaluación de la investigación del profesorado universitario español: métricas y productos*. Mesa redonda: políticas de evaluación del profesorado. UNIR, 11 de diciembre de 2019.

Delgado-López-Cózar, Emilio; Martín-Martín, Alberto (2019). "El factor de impacto de las revistas científicas sigue siendo ese número que devora la ciencia española: ¿hasta cuándo?". *Anuario ThinkEPI*, v. 13, e13e09. <https://doi.org/10.3145/thinkepi.2019.e13e09>

Dialnet Métricas (2019). *Índice Dialnet de revistas e investigadores*. <https://dialnet.unirioja.es/metricas/>

Google Scholar (2021). *Inclusion guidelines for webmasters*. <https://scholar.google.es/intl/es/scholar/inclusion.html>

Ministerio de Ciencia e Innovación (2021). "La Agencia Estatal de Investigación se adhiere a la Declaración de San Francisco sobre la evaluación científica (DORA)". *Ministerio de Ciencia e Innovación*, 20 abril. <https://cutt.ly/1ngZ7Mi>

Antonio Calderón-Rehecho
Universidad Complutense de Madrid
acaldero@ucm.es

* * *

No deben utilizarse rankings de revistas para la evaluación de individuos

Isidro F. Aguillo

Al hilo de la nota de Andoni, comentaré un par de ideas de forma somera:



1) No tiene sentido seguir desarrollando y usando indicadores de revistas cuyo fin último es la evaluación de individuos. España es, con diferencia, el mayor productor del mundo de rankings de revistas (*MIAR*, *Dialnet Métricas*, *Carhus*, *SJR*, *Fecyt*, *InReX*, *CIRC* y otras) en un momento en que las métricas de artículos, por muchas razones, son una solución mucho mejor y totalmente viable y fiable para los propósitos de la evaluación.

2) Sin entrar a discutir los aberrantes cuartiles, un sinsentido bien descrito por Andoni, me gustaría centrarme en las fuentes y sus indicadores. Todas las métricas (relevantes) de revistas están basadas en el análisis de citas. Y hay tres grandes grupos de citas:

- citas esperadas: lo que es en realidad el factor de impacto, que por cierto debería ser un número entero;
- citas recogidas: citas reales extraídas de los universos de cobertura de las diferentes fuentes;
- citas totales: todas las citas reales recibidas por un trabajo publicado.

Podemos debatir el valor de las citas y el sentido de cada cita en particular, pero ignorar una fracción significativa de ellas es fundamentalmente un grave error. Comparemos las coberturas (con métricas) de revistas de diferentes fuentes:

- *Dialnet Métricas*: unos pocos miles;
- *WoS/JCR*: unas 12.000;
- *Scopus*: unas 25.000.

El número total de revistas con revisión por pares supera las 40.000 (*MIAR*) y posiblemente sean más de 60.000, todo ello sin contar otras muchas fuentes de citas (libros, congresos, patentes, tesis, ...). Entonces, ¿qué sentido tiene utilizar citas esperadas de menos de cuatro mil revistas Q1, cuando podemos tener información de citas completas de un universo 20 veces mayor? Sin duda, en ese escenario no habría preocupación alguna por la tercera cifra decimal.

Isidro F. Aguillo
CSIC
isidro.aguillo@cchs.csic.es

* * *

Las aberraciones de los índices de impacto de revistas: una vez más... como un disco rayado

Emilio Delgado-López-Cózar



Dos son, entre otras, las cuestiones planteadas en el debate suscitado por Andoni:

- El sentido, corrección y validez del factor de impacto de una revista, y por afinidad, del resto de sucedáneos bibliométricos: hermanos, primos hermanos o primos a secas.
- El uso, abuso, mal uso de estas métricas aplicadas a la evaluación individual de los investigadores, y en el caso de España, a todo lo que se mueva en la escena académica: instituciones, programas y un largo etcétera...

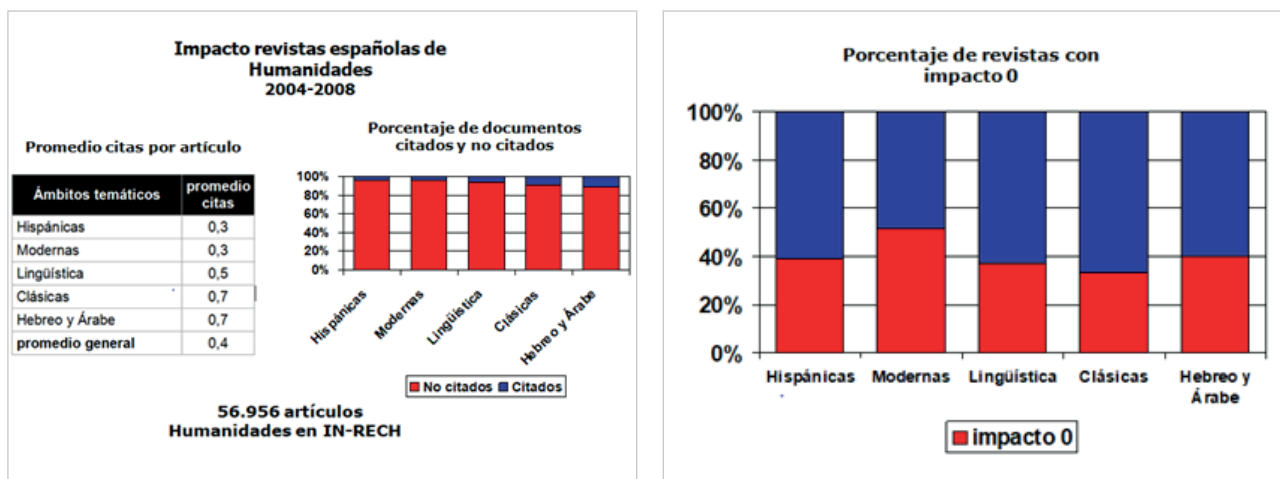
Respecto al primer punto, la nota de Andoni, que confiesa hecha desde "la realidad", yo diría que con un lenguaje llano, lo que evidencia, una vez más, son los desvaríos "métricos" de los índices de impacto. Digo una vez más, porque como apuntaba Isidro son muchos y variados los trabajos que arrojan evidencias empíricas sobre ello.

Lo intentaré expresar de una manera lo más simple y llana posible:

- ¿Qué significa poseer un impacto de 200 centésimas de cita en promedio en cinco años para una revista que ha publicado 150 artículos?
- ¿Qué diferencia existe respecto a una revista que tenga 800 centésimas de cita, que pasaría por ser en todas las disciplinas como una revista TOP?
- Por favor, reparad en el absurdo ¿ni siquiera una cita en cinco años? Y, por cierto, ¿es que las citas se pueden cuartear como los solomillos? "¡¡Deme 150 gramos de ternera...!!" Qué feliz y contento este autor español que proclama sin rubor "¡¡¡Yo tengo 250 centésimas de cita porque mi artículo se ha publicado en una revista del cuartil 1 que tiene un índice de impacto de 0,250...; con un par más de estos ya tengo un sexenio y engroso la cuenta de resultados para mi acreditación en la *Aneca*... !!!"

Por si no fuera ya suficiente, resulta que aquí se cumplen las famosas leyes bibliométricas que explican la concentración y/o dispersión de los fenómenos bibliográficos. Siguiendo con el estilo simple y sencillo recurramos al principio de Pareto (80/20) que subyace a toda distribución bibliométrica. Solo un 20% de los trabajos concentrará el 80% de las citas, mientras que el 80% de trabajos concitará el 20% de citas restante. Es algo que ya sabemos desde hace muchos años: la inmensa mayoría de los trabajos no son citados, sobre todo, en Humanidades y muchas Ciencias Sociales.

Permitidme que visualice el fenómeno con unas viejas gráficas que ilustraban lo acaecido en aquel experimento que fue IN-RECH. Y como las gráficas hablan por sí solas, les ahorro más monsergas. Sólo diré que son muy pocos los artículos y los autores que podrían presumir del impacto de las revistas donde han publicado, o lo que es lo mismo, es absolutamente falaz atribuirse el impacto de la revista como propio, además de grotesco con números tan pírricos.



Pues bien, dicho esto me da igual:

- El indicador que empleemos: ¿Qué significa tener un índice h de 6? ¿6 artículos con al menos 6 citas por los más de 100 artículos publicados en cinco años!!! y si el índice h es de 15 ¿hay diferencia alguna en su insignificante y grotesco índice h!!!
- Que se midan citas a secas o con pesos, atendiendo al contexto...
- Que lo hagan con unas bases de datos u otras: *Dialnet*, *Web of Science* o *Scopus*. Números algo más grandes, ¿y qué? ¿acaso cambia la naturaleza o la esencia del problema?

Al final, llega la incómoda pregunta:

¿Cómo estos números se pueden emplear para valorar positivamente el curriculum de un investigador como se hace en España y en muchos lugares del mundo?

Y con este interrogante me adentro en el segundo punto: el uso, abuso y mal uso del indicador. Y aquí nada más esclarecedor que recurrir a la *Declaración de San Francisco sobre la Evaluación de la Investigación*, conocida popularmente como *DORA*. Entresaco algunas de las recomendaciones más destacadas en línea con lo que venimos tratando:

- No utilice métricas basadas en revistas, como el factor de impacto, como una medida sustituta de la calidad de los artículos de investigación individuales, para evaluar las contribuciones de un científico individual, o en las decisiones de contratación, promoción o financiación.
- El contenido científico de un artículo es mucho más importante que las métricas de publicación o la identidad de la revista en la que fue publicado.
- Considere una amplia gama de medidas de impacto que incluyan indicadores cualitativos, como la influencia sobre la política y prácticas científicas.
- Reduzca profundamente el énfasis en el factor de impacto como herramienta promocional, idealmente dejando de promover su uso o presentando la métrica en el contexto de una variedad de métricas basadas en revistas.
- Ponga a disposición una variedad de métricas a nivel de artículo para alentar un cambio hacia la evaluación basada en el contenido científico de un artículo en lugar de las métricas de publicación de la revista en la que se publicó.
- Cuando participe en comités que toman decisiones sobre financiación, contratación, permanencia o promoción, realice evaluaciones basadas en el contenido científico en lugar de en métricas de publicación.
- Impugne las prácticas de evaluación que dependen indebidamente del factor de impacto y promueva y enseñe prácticas que se centren en el valor y la influencia de los resultados de investigación específicos.

Llegados hasta aquí y a la vista de la preeminencia absoluta del factor de impacto en la política científica española reproduciré las preguntas que ya planteábamos en una nota *Thinkepi* que dedicábamos monográficamente al tema (Delgado-López-Cózar; Martín-Martín, 2019):

- ¿Cómo es posible que el FI siga reinando en España cual monarca absoluto, pese a todo y a todos?
- ¿Es que los científicos, gestores españoles y responsables de políticas científicas españolas no saben leer o es no quieren leer estas declaraciones? OJO: la *Agencia Española de Investigación* firmó la declaración. ¿Será un antes y un después de esta decisión en la evaluación científica en España?
- ¿Cómo es posible que científicos de todas las especialidades, sesudos en sus campos de conocimiento,

que aplican complejas herramientas, rigurosos procedimientos metodológicos y sofisticadas técnicas de análisis de datos se dejen seducir por un simple número, un promedio que es la medida estadística más elemental que existe, para medir el éxito de todo su trabajo y trayectoria?

- ¿Cómo es posible que acepten juzgar su trabajo por el envoltorio en el que se presenta y vende y no atiendan a su valor singular, real y contrastado? ¿Es que no saben distinguir el medio del mensaje, el recipiente o continente del contenido?
- ¿Cómo es posible que un científico que en su trabajo diario no prejuzga nada, no admite más que evidencias basadas en hechos no en presunciones, que considera que un trabajo no necesariamente es bueno porque haya sido firmado por un prestigioso autor de una prestigiosa universidad, ni malo porque haya sido realizado por autores desconocidos de mediocres instituciones, dé por sentado que es excelente por ser publicado en una revista de impacto?
- ¿Por qué se abraza el FI casi como objeto de culto, un auténtico tótem?

¿Es que no hay alternativas? Las hay tanto para el gusto de impactólatras como impactófobos. Dejo tinta en el tintero...

Delgado-López-Cózar, Emilio; Martín-Martín, Alberto (2019). "El factor de impacto de las revistas científicas sigue siendo ese número que devora la ciencia española: ¿hasta cuándo?". *Anuario ThinkEPI*, v. 13, e13e09. <https://doi.org/10.3145/thinkepi.2019.e13e09>

Emilio Delgado-López-Cózar
Universidad de Granada
edelgado@ugr.es

* * *

Por un cambio radical en la evaluación de la investigación en España

Emilio Delgado-López-Cózar

Cuando el otro día me despedía de este foro diciendo que me dejaba tinta en el tintero, me refería literalmente al hecho de que tres colegas estábamos a punto de publicar un llamamiento a las autoridades españolas para cambiar el modelo de evaluación de la investigación en España, que tanta importancia concede a los indicadores bibliométricos, en general, y a los índices de impacto de revista, en particular.

Pues bien, dicho llamamiento firmado por Ernest Abadal, Ismael Ràfols y yo mismo, se ha publicado en forma de carta en *Profesional de la Información* (**Delgado-López-Cózar; Ràfols; Abadal**, 2021).

Hace unas semanas la *Agencia Española de Investigación* se adhirió a la *Declaración DORA* sobre evaluación de la investigación. Promete "revisar todas sus convocatorias y documentos de soporte a la evaluación para asegurar su correcta implementación".

A pesar del gran paso dado por esta agencia no será fácil cambiar políticas que por su larga permanencia en el tiempo han penetrado en lo más hondo del tejido científico español.

Los profesionales de las bibliotecas, documentación e información sois elementos clave para cambiar esta cultura. Os pedimos ayuda para que diseminéis este documento entre profesores e investigadores de universidades y centros de investigación, donde el tema, sabéis muy bien y de primera mano, despierta todo tipo de emociones por cuanto afecta decisivamente a los propios pilares del quehacer cotidiano de los científicos.

Estamos pensando en otras iniciativas que abran un debate nacional profundo y sin restricciones. Os mantendremos informados.

Gracias por vuestra atención y apoyo,

Delgado-López-Cózar, Emilio; Ràfols, Ismael; Abadal, Ernest (2021). "Letter: A call for a radical change in research evaluation in Spain". *Profesional de la información*, v. 30, n. 3, e300309. <https://doi.org/10.3145/epi.2021.may.09>

Emilio Delgado-López-Cózar
Universidad de Granada
edelgado@ugr.es