

Bibliometría e indicadores de actividad científica (III). Indicadores de impacto basados en las citas (1)

R. Aleixandre-Benavent^{1,2}, J. González de Dios³, L. Castelló Cogollos^{1,4}, C. Navarro Molina^{1,5},
A. Alonso-Arroyo^{1,5}, A. Vidal-Infer^{1,5}, R. Lucas-Domínguez^{1,5}

¹Unidad de Información e Investigación Social y Sanitaria-UISYS (CSIC-Universitat de València). Valencia. ²Instituto de Gestión de la Innovación y del Conocimiento-Ingenio (CSIC-Universitat Politècnica de València). ³Servicio de Pediatría. Hospital General Universitario de Alicante. Departamento de Pediatría. Universidad «Miguel Hernández». Alicante.

⁴Departamento de Sociología y Antropología Social. ⁵Departamento de Historia de la Ciencia y Documentación. Universitat de València

Resumen

Los indicadores bibliométricos basados en las citas se apoyan en el supuesto de que los trabajos importantes son usualmente citados, mientras que los irrelevantes se ignoran. En esta entrega se exponen los supuestos básicos del análisis de citas y el método de cálculo de los indicadores, así como sus ventajas, inconvenientes y limitaciones. También se discuten las estrategias en las que se apoyan las revistas y los investigadores para aumentar el número de citaciones y los artículos más citados en la revista *Acta Pediátrica Española*. Los indicadores básicos más utilizados son el número de citas recibidas por una unidad de investigación (autor, grupo, institución, país, revista o área temática), el promedio de citas por trabajo, el número de trabajos muy citados, el índice h, los indicadores basados en el índice h y el indicador corona (*crown indicator*). Entre los indicadores extraídos del análisis de referencias se exponen la vida media, el índice de Price y el índice de aislamiento.

Palabras clave

Bibliometría, indicadores bibliométricos, citas, vida media, artículos más citados, índice h, índice de Price, indicador corona

Abstract

Title: Bibliometrics and indicators of scientific activity (III). Citation impact based indicators (1)

The bibliometric indicators based on the citations are founded on the assumption that the important papers are usually cited, while the irrelevant ones are ignored. The basic assumptions of citation analysis, the method of calculating these indicators, their advantages, drawbacks and limitations are discussed. We also outline strategies in which journal editors and researchers are supported to increase the number of citations and the most cited articles in the journal *Acta Pediátrica Española*. The most commonly used basic indicators are the number of citations received by a research unit (author, group, institution, country, journal or subject area), average citations per paper, number of papers cited, h-index, indicators based on the h-index and the crown indicator. Among the indicators extracted from the analysis of references, we discuss the cited half-life, the Price index and the isolation index.

Keywords

Bibliometry, bibliometric indicators, citations, half-life, most cited articles, h-index, Price index, crown indicator

Citas, referencias y supuestos generales del análisis de citas

La moderna normativa científica requiere que, cuando un investigador publica un trabajo, se refiera a trabajos anteriores relacionados con el tema y los cite. Estas referencias conducen a las fuentes de las ideas contenidas en el trabajo. Por tanto, podemos distinguir entre «citas» (las que una publicación recibe de trabajos posteriores) y «referencias» (las que una publicación hace a otras anteriores)¹⁻³.

Los conceptos de cita y referencia no son intercambiables. Como se aprecia en la figura 1, si un artículo C contiene una nota bibliográfica en que se describe y se utiliza otro artículo

A, se dice que C contiene una referencia a A y que A recibe una cita de C. El número de referencias contenidas en un artículo se determina por el número de trabajos mencionados en su bibliografía, mientras que el número de citas que recibe un artículo se halla mediante la consulta de algún índice que informe de cuántos artículos lo mencionan (un índice de citas)³. En la figura 2 se muestra, a modo de ejemplo, un flujo entre trabajos citadores y citados: los trabajos A, B y C han recibido 8 citas procedentes de 6 trabajos citadores y una de ellas es una autocita. Hasta hace relativamente pocos años sólo existían dos fuentes en las que se podía conocer el número de citas que había recibido un investigador o una revista, el Science Citation Index y el Social Science Citation Index del Institute for Scientific Information, fundado por Eugene Garfield en

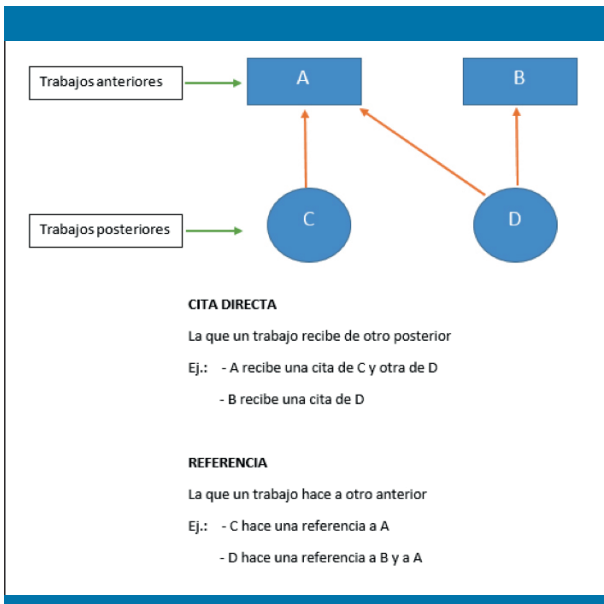


Figura 1. Concepto de cita y referencia

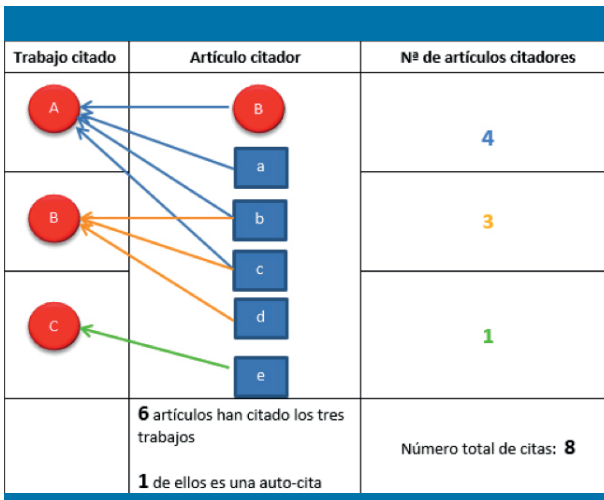


Figura 2. Flujo entre trabajos citadores y citados

1960, posteriormente adquirido por Thomson Reuters y propiedad de Clarivate Analytics desde 2016. Sin embargo, hoy en día disponemos de varias bases de datos de citas, como Scopus, Google Scholar y SciELO Citation Index, esta última incluida en la plataforma Web of Science, al igual que Science Citation Index y Social Science Citation Index⁴⁻⁶.

Los indicadores procedentes del análisis de citas y referencias se apoyan en el supuesto de que los trabajos importantes son usualmente citados, mientras que los irrelevantes se ignoran. Sin embargo, la investigación sociológica en torno a las citas y referencias ha demostrado que la realidad es mucho más compleja.

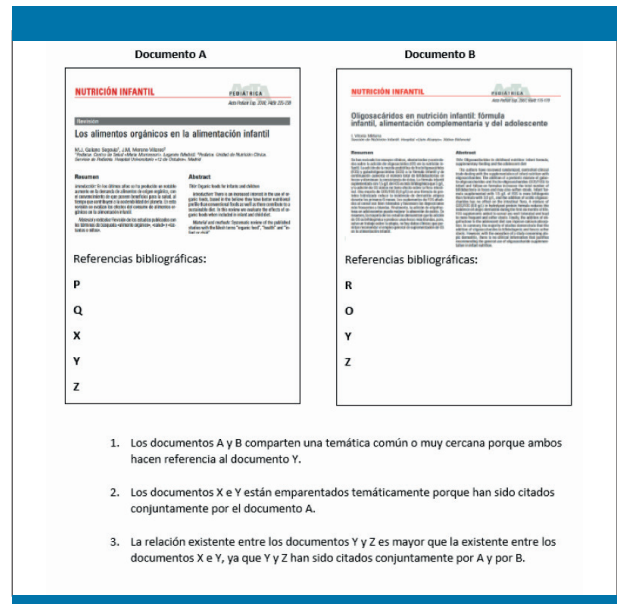


Figura 3. Relaciones que se establecen entre documentos a partir de las citas

El análisis de citas parte de los siguientes supuestos generales^{2,3,7,8}:

- Un documento hace referencia a un trabajo anterior si éste aporta algún tipo de información y ambos documentos tienen una temática común.
- Si dos documentos hacen referencia a un mismo trabajo, comparten una temática común o muy cercana.
- Si dos trabajos son citados conjuntamente por dos o más documentos, están emparentados temáticamente.
- Los documentos de un mismo autor tienden a tratar idénticos o similares temas.

En la figura 3, en la que se presentan dos documentos A y B con sus referencias bibliográficas, se ilustran gráficamente las afirmaciones anteriores.

Indicadores bibliométricos básicos basados en las citas

Se trata de un grupo de indicadores muy importante, ya que la mayoría de los indicadores bibliométricos propuestos en la literatura son variantes de éstos. El número de publicaciones de una unidad de investigación (p. ej., un investigador, un grupo de investigación, un país o una revista) y el número de citas de estas publicaciones es probable que no coincidan en diferentes bases de datos, debido a que la cobertura de revistas es variable en cada una. El número de publicaciones y de citas de una unidad de investigación también depende del periodo considerado dentro del cual se cuentan las publicaciones y las citas.

Una decisión muy importante que debe tomarse en los recuentos de citas es la selección de la «ventana de citación», es decir, el número de años retrospectivos que deben considerarse para cuantificar las citas. Algunos estudios han concluido que la elección de 2 años, utilizada en el cálculo del factor de impacto, puede ser suficiente en ciertas áreas, como las biomédicas⁹, la física y algunas incluidas en ciencias de la vida¹⁰. Sin embargo, en otras áreas que tienen una dinámica de citación más lenta y, por tanto, necesitan más tiempo para recibir citas, este periodo es demasiado breve y debería ampliarse para que la mayoría de las publicaciones puedan ser reconocidas y citadas^{9,11-13}. En esta línea, el Journal Citation Reports de Web of Science introdujo en 2007 un nuevo indicador para sus revistas, el factor de impacto de 5 años, que, como se verá en el próximo capítulo, contempla una ventana de citación de 5 años y permite complementar el factor de impacto de 2 años¹⁴.

Pueden distinguirse varios indicadores básicos de impacto de citas: número total de citas, número promedio de citas por publicación, número de publicaciones muy citadas, porcentaje de publicaciones altamente citadas, índice h, indicadores de la familia h e indicador corona (*crown indicator*)^{6,7,15,16}.

Número total de citas

Consiste en contar el número total de citas de las publicaciones de una unidad de investigación, que, como ya hemos comentado, puede ser un investigador, un grupo, una institución o un país. Por ejemplo, si los trabajos de una unidad de investigación han recibido 60, 30, 25, 20, 20, 15, 10, 10, 5, 2, 1 y 0 citas, el número total de citas es de 198.

Si el número de trabajos que publica un investigador es muy importante para su carrera, el número de citas también lo es porque indica el impacto que su investigación tiene en el campo. Las agencias de financiación a menudo buscan una combinación de ambos parámetros para tomar decisiones. Sin embargo, el número de citas que recibe un trabajo no es, en principio, una medida de su calidad científica, aspecto que sólo podrá ser valorado por los expertos en el área³. Es difícil predecir en la citación qué proporción se debe a la calidad intrínseca del trabajo citado y cuánta a otros factores, como el prestigio de la revista citada, el prestigio de la institución a la que pertenece el autor, etc., ya que el fenómeno de la citación está sujeto a modas, fobias y otras tendencias.

Las razones por las que un autor puede citar a otro son muy variadas³:

1. Rendir homenaje a los pioneros de la disciplina de estudio.
2. Dar crédito al trabajo referido (homenaje a los iguales).
3. Identificar la metodología, el equipo, los instrumentos de medida, las variables, etc.
4. Suministrar una bibliografía de base.
5. Corregir el propio trabajo o el trabajo de otros.
6. Criticar trabajos previos.
7. Alertar o advertir sobre circunstancias que tener en cuenta en próximos trabajos.

8. Suministrar información sobre trabajos poco difundidos, no indizados o poco citados.
9. Autentificar datos o hechos, como constantes físicas, datos de laboratorio de otros trabajos, etc.
10. Identificar publicaciones originales que describían un concepto o término epónimo.
11. Rechazar ideas de otros o las prioridades defendidas por otros (homenaje negativo).

Aunque esta lista proporciona una buena indicación de las razones que pueden llevar a un autor a citar a otro, es difícil asegurar exactamente cuáles son las que le llevan a hacerlo en cada caso; pero lo que sí es seguro es que no se trata únicamente de definir las bases de la idea científica propuesta en su trabajo. Para analizar el alcance y la significación de una cita determinada, habría que evaluar el contexto de la cita, es decir, la frase o las frases que en el texto acompañan a la cita. Así podremos determinar si es positiva (los resultados del artículo citado son admitidos sin discusión) o si su presencia obedece simplemente a cuestiones de forma³.

Aunque el número de citas que recibe un trabajo no es una medida de su calidad científica, sino indicativo de su visibilidad, uso, difusión o impacto, según varios autores existe una correlación positiva entre la clasificación de trabajos en función de las citas recibidas y el juicio de los expertos u otros indicadores de calidad de la investigación.

Otros factores que afectan a la función normal de las citas son los siguientes³:

- Cuanto más general es el tema que se trata, más posibilidades tiene de ser citado.
- Los autores o publicaciones extraordinariamente importantes tienden a darse por conocidos y no ser citados explícitamente (fenómeno de la obliteración).
- Algunos trabajos se incorporan tan rápidamente al patrimonio del saber colectivo que ya nadie manifiesta ninguna necesidad de citarlos porque son ampliamente conocidos.
- Existe un elevado número de citas perfunctorias, es decir, realizadas a la ligera, por formulismo o para salvar las apariencias.
- Existen barreras idiomáticas y nacionales que dificultan la citación.
- Se producen falsificaciones y manipulaciones de las citas.

Número promedio de citas por publicación

Consiste en hallar el número medio de citas de las publicaciones de una unidad de investigación. Para la unidad de investigación del ejemplo anterior, el número medio de citas por publicación es de $198/12 = 16,5$ citas por trabajo. El indicador más conocido, basado en contar el promedio de citas recibidas por los artículos de una revista, es el factor de impacto, que se describirá y analizará en el próximo capítulo. Los indicadores basados en los promedios de citas son muy criticados debido a que las distribuciones de citas no son homogéneas, sino que están desviadas por la existencia tanto de artículos muy cita-

TABLA 1

Artículos más citados de *Acta Pediátrica Española* en Scopus, y Ncc (8 o más)

Autores	Título	Fuente	N.º de citas	N.º de años hasta 2016	Ncc
Moreno Villares JM, Galiano Segovia MJ	La comida en familia: algo más que comer juntos	Acta Pediatr Esp. 2006; 64: 554-558	12	10	1,2
Galiano Segovia MJ, Moreno Villares JM	El desayuno en la infancia: más que una buena costumbre	Acta Pediatr Esp. 2010; 68: 403-408	12	6	2
González de Dios J, Buñuel Álvarez JC, González Rodríguez R, Alonso Arroyo A, Aleixandre Benavent R	Fuentes de información bibliográfica (XIV). Sobre «fuentes», «pirámides» y «revoluciones» en la gestión del conocimiento en pediatría	Acta Pediatr Esp. 2012; 70: 289-295	10	4	2,5
Cabo Masip T, Alentado Morell N, Dalmau Serra J	Nuevas recomendaciones diarias de ingesta de calcio y vitamina D: prevención del raquitismo nutricional	Acta Pediatr Esp. 2008; 66: 233-236	9	8	1,1
Coronado Ferrer S, Peset Mancebo F, Ferrer Sapena A, González de Dios J, Aleixandre-Benavent R	Web 2.0 en medicina y pediatría (I)	Acta Pediatr Esp. 2011; 69: 3-11	9	5	1,8
Aleixandre-Benavent R, González Alcaide G, González de Dios J, Alonso-Arroyo A	Fuentes de información bibliográfica (I). Fundamentos para la realización de búsquedas bibliográficas	Acta Pediatr Esp. 2011; 69: 131-136	9	5	1,8
Moreno Villares JM, Oliveros Leal L, Pedrón Giner C	Desnutrición hospitalaria en niños	Acta Pediatr Esp. 2005; 63: 63-69	9	11	0,8
González Herrero M, Toledano Bueno J	La lactancia materna en nuestro medio: análisis de la situación	Acta Pediatr Esp. 2007; 65: 123-125	8	9	0,9
Sanz Y, Collado MC, Dalmau J	Probióticos: criterios de calidad y orientaciones para el consumo	Acta Pediatr Esp. 2003; 61: 476-482	8	13	0,6
Vitoria Miñana I	Oligosacáridos en nutrición infantil: fórmula infantil, alimentación complementaria y del adolescente	Acta Pediatr Esp. 2007; 65: 175-179	8	9	0,9
Balaguer Santamaría A, González de Dios J	Digiriendo la información científica. Excelencias y limitaciones de las revisiones sistemáticas y metaanálisis	Acta Pediatr Esp. 2004; 62: 4-10	8	12	0,66
Durá Travé T, Sánchez-Valverde Visus F	Obesidad infantil: ¿un problema de educación individual, familiar o social?	Acta Pediatr Esp. 2005; 63: 204-207	8	11	0,7
Sanz Y, Collado MC, Haros M, Dalmau J	Funciones metabólico-nutritivas de la microbiota intestinal y su modulación a través de la dieta: probióticos y prebióticos	Acta Pediatr Esp. 2004; 62: 520-526	8	12	0,66

Ncc: número de citas corregido, resultado de dividir el número de citas recibido por el número de años transcurridos desde la publicación del artículo hasta 2016.

dos como de otros que no son apenas citados o no se citan nunca¹⁵⁻¹⁸.

Número de publicaciones muy citadas

Consiste en determinar el número de publicaciones de una unidad de investigación que se consideran altamente citadas aplicando un umbral determinado. Si en el ejemplo anterior utilizamos un umbral de más de 20 citas, la unidad de investigación tiene 3 publicaciones muy citadas. Algunos autores consideran las publicaciones muy citadas como indicios de excelencia científica^{19,20}. El índice i10, reportado por Google Scholar, se basa en la idea de contar publicaciones altamente citadas, ya que recoge las publicaciones que se han citado al menos 10 veces.

En la tabla 1 se presentan los artículos de *Acta Pediátrica Española* recogidos en la base de datos Scopus que han sido cita-

dos 8 o más veces. Dos artículos han recibido 12 citas: el publicado en 2006 por Moreno Villares y Galiano Segovia, titulado «La comida en familia: algo más que comer juntos», y el publicado en 2010 por estos mismo autores, titulado «El desayuno en la infancia: más que una buena costumbre». Les sigue el artículo publicado en 2012 por González de Dios et al., titulado «Fuentes de información bibliográfica (XIV). Sobre “fuentes”, “pirámides” y “revoluciones” en la gestión del conocimiento en pediatría». Debido a que las publicaciones más antiguas han tenido más posibilidades de ser citadas que las más actuales, en ocasiones es necesario establecer un corrector.

Una variante correctora de este índice consiste en dividir el número de citas por los años transcurridos desde la publicación del artículo hasta la actualidad. Si en la tabla anterior aplicamos este corrector considerando los años transcurridos hasta 2016,

el artículo con un mayor número de citas corregido (Ncc) es ahora el de González de Dios et al. mencionado antes (Ncc= 2,5), seguido del artículo de Moreno Villares y Galiano Segovia de 2010 (Ncc= 2) y de dos trabajos con un Ncc de 1,8: el publicado por Coronado Ferrer et al. en 2011, titulado «Web 2.0 en medicina y pediatría (I)», y el publicado por Aleixandre-Benavent et al. en 2011, titulado «Fuentes de información bibliográfica (I). Fundamentos para la realización de búsquedas bibliográficas».

Porcentaje de publicaciones altamente citadas

Consiste en determinar el porcentaje de trabajos muy citados respecto del total de trabajos. En nuestro ejemplo, el porcentaje es de $3/12= 0,25$ (es decir, el 25%).

Índice h (o índice de Hirsch)

El índice h es un indicador desarrollado por Jorge Hirsch en 2005, físico teórico argentino de la Universidad de California, ante la necesidad de medir o evaluar la investigación individual. Posteriormente, se amplió su utilidad a la valoración de revistas. Según Hirsch²¹, un investigador tiene un índice h determinado cuando *h* de sus artículos han recibido al menos *h* citas cada uno y las otras publicaciones tienen menos de *h* citas. Es decir, un índice h de 30 significa que la unidad de investigación tiene 30 artículos con al menos 30 citas cada uno. Para calcularlo, se ordenan los trabajos por número de citas de mayor a menor, y el índice h es un guarismo que expresa la coincidencia del número de artículos publicados con el número de citaciones recibidas²².

Para la unidad de investigación en nuestro ejemplo, el índice h es de 8, ya que hay 8 publicaciones cada una de las cuales tiene al menos 8 citas, mientras que las otras 4 publicaciones tienen cada una menos de 8 citas.

El índice h ha motivado muchos estudios que analizan sus ventajas y limitaciones, así como la propuesta de indicadores complementarios²³⁻²⁸. La principal ventaja del índice h, aparte de la facilidad de obtención, es que combina en un solo indicador una medida de cantidad (número de publicaciones) y otra de la calidad (número de citas recibidas). Por tanto, es un indicador útil para detectar a los investigadores más destacados en un área²³. Otra ventaja reside en que mide las citas de trabajos concretos, sin tener en cuenta revistas determinadas, y en su cálculo todas las citas tienen el mismo valor. Además, se puede aplicar en cualquier periodo de tiempo.

Las principales limitaciones del índice h son^{15,22-24}:

1. Es inadecuado para comparar investigadores de diferentes áreas científicas, lo que se explica por los distintos hábitos de publicación y citación según el campo. Un ejemplo de ello es que los investigadores más relevantes en biología doblan en sus valores de índice h a los de otras áreas, como la física, y éstos, a su vez, se sitúan muy por encima de las matemáticas.
2. Muestra una correlación positiva con el número total de citas y documentos de los investigadores, por lo que tiende a favorecer a los que tienen carreras científicas más largas, y tiene menos validez en los más jóvenes con un bajo número

de publicaciones. La dependencia entre el índice h de un científico y su número total de documentos hace que el indicador nunca pueda ser superior al número total de documentos. De este modo, un investigador que haya publicado 30 trabajos nunca podrá tener un índice h superior a 30, independientemente del número de citas que hayan recibido sus trabajos. Por otra parte, el número total de citas tampoco influye en el valor de *h*, de manera que, por ejemplo, un investigador que tenga 30 artículos con 30 citas cada uno (900 citas en total) tendría un índice h de 30, mientras que otro investigador que tenga 30 artículos con 40 citas cada uno (1.200 citas en total) tendría el mismo índice, y un investigador con 20 artículos con 70 citas cada uno (1.400 citas en total) solamente tendría un índice h de 20. En otras palabras, los investigadores con pocas pero muy relevantes publicaciones siempre tendrán un índice h bajo.

3. No tiene en cuenta la calidad de las revistas en las que un autor publica sus investigaciones, lo que constituye una limitación importante, dado que existen notables diferencias entre publicaciones en cuanto a los filtros de calidad que imponen a los trabajos que publican. Por tanto, el valor del índice h puede verse artificialmente implementado por citas de revistas de calidad media o baja. Puede darse la paradoja de que un trabajo sea muy citado por determinados colegas, aunque su calidad no sea excelente.
4. Puede incrementarse artificialmente si los investigadores citan exageradamente sus documentos más próximos a entrar en el cálculo del índice h. A pesar de que la investigación es un proceso acumulativo en el que es normal un cierto grado de autocitación, algunos autores sugieren la exclusión de las autocitas.
5. No es sensible a las variaciones del rendimiento de personas o instituciones en periodos de tiempo diferentes. Consecuentemente, los investigadores que cesen en su actividad científica mantienen constante su índice h, aunque no vuelvan a publicar ningún otro artículo. Es decir, el índice h no es capaz de diferenciar entre investigadores activos e inactivos²².
6. Está sujeto a las mismas limitaciones que tienen las bases de datos que incluyen citas, debido a los problemas de homonimia y a la falta de normalización de las variantes de firma de los autores. Su valor es diferente según la base de datos utilizada para su cálculo (Web of Science, Scopus, Google Scholar), por lo que algunos autores recomiendan obtenerlo partiendo de una relación completa de las publicaciones de los investigadores, validada por ellos mismos, para aumentar la fiabilidad de los resultados.
7. Por último, no parece adecuado evaluar a un científico o grupo con un solo dígito, ya que la investigación científica es muy compleja y multidimensional y no se puede evaluar con un indicador simple, sino que requiere combinar varios indicadores que determinen la producción y el impacto científico de un autor o institución. Otra limitación es que el índice h aporta poca información, más allá del número total de citaciones, por lo que no sería correcto considerarlo como la medida global de la calidad de la investigación de un autor o institución.

TABLA 2	Indicadores básicos basados en las citas	
	Indicadores dependientes del tamaño	Indicadores independientes del tamaño
	Número total de citas	Número medio de citas por publicación
	Número de publicaciones muy citadas	Porcentaje de publicaciones altamente citadas
	Índice h	

En la tabla 2 se clasifican los cinco indicadores anteriores según sean o no dependientes del tamaño de la muestra de publicaciones. Los indicadores dependientes del tamaño no disminuyen nunca cuando se obtienen publicaciones adicionales y permiten disponer de una medida global del impacto de la unidad de investigación. Por el contrario, los indicadores independientes del tamaño pretenden proporcionar una medida de rendimiento promedio por publicación y pueden disminuir cuando se obtienen publicaciones adicionales. Normalmente se utilizan indicadores independientes del tamaño para hacer comparaciones entre unidades de diferente tamaño (p. ej., entre un grupo de investigación pequeño y uno grande, o entre una universidad pequeña y una universidad grande)¹⁵.

Indicadores de la familia h: índices g, A, R y AR

En 2006 fue propuesto el índice g, que puede alcanzar valores superiores al número total de documentos de un autor, y tiene en cuenta la cantidad de citas recibidas por los documentos incluidos en su cálculo^{29,30}. A igual que el índice h, uno de sus inconvenientes es que puede verse excesivamente influido por publicaciones ocasionales muy citadas, no representativas de la actividad científica del investigador²³.

Algunas propuestas recientes sugieren complementar el índice h con otros índices que tienen en cuenta las citas recibidas por los documentos del denominado *h-core* (los que determinan el índice h), lo que podría beneficiar a los autores selectivos. Así, se ha descrito el índice A, que calcula el número medio de citas recibidas por los artículos del *h-core*; el índice R, que mide la intensidad de citación de dicho núcleo, y el índice AR, que tiene en cuenta el año de publicación de los documentos. No obstante, no son los sustitutos del índice h, sino que deben utilizarse en combinación con él para superar algunas de sus limitaciones^{23,31-34}.

«Crown indicator»

El indicador corona, propuesto por el Center for Science and Technologies Studies de la Universidad de Leiden, mide el impacto científico de un investigador o un grupo de investigación. Se calcula dividiendo el número medio de citas recibidas por documento del investigador o grupo (excluyendo las autocitas) por el número medio de citaciones de todos los artículos de un determinado campo científico. El indicador se expresa como un número decimal que muestra la relación con el promedio, que

es 1. A partir de aquí, se puede concluir que la calidad de un grupo científico está muy por debajo (<0,5), por debajo (0,5-0,8), en torno a (0,8-1,2), por encima (1,2-2) o muy por encima (>2) del nivel internacional. Al tener en cuenta la tasa de citación del campo de investigación, el indicador corona supera las limitaciones de otros indicadores, debidas a los hábitos de citación y dinámica de las citas, así como el peso que tienen determinados artículos (como los de revisión), ya que se citan más frecuentemente que los originales^{16,35,36}.

Indicadores basados en el análisis de referencias

Obsolescencia y semiperiodo (*vida media*)

Un problema inseparable del crecimiento exponencial de la ciencia moderna es el rápido envejecimiento de la literatura científica, es decir, la acelerada tendencia a que las publicaciones científicas caigan en desuso (obsolescencia). Un documento científico «envejece» a medida que nuevos trabajos incorporan datos más recientes al frente de investigación abierto. Además, cuanto mayor es el grado de actividad en un área determinada, mayor es el número de trabajos de reciente publicación que suceden a los anteriores y, por tanto, disminuye la posibilidad de que un trabajo antiguo sea citado³.

En 1960, Burton y Kebler³⁷ proporcionaron el indicador de dicho proceso con su concepto de semiperiodo, o semivida (*half-life*), de la literatura científica. Tomando la definición de semiperiodo de desintegración de la física nuclear (tiempo requerido para que se produzca la desintegración de la mitad de los átomos de una sustancia radiactiva), aplicaron este concepto a la literatura científica como «el tiempo durante el cual fue publicada la mitad de la literatura activa circulante sobre un tema determinado». Estos autores basaron la técnica de su determinación en el análisis de las referencias bibliográficas distribuidas por años de procedencia y consideraron las publicaciones citadas en un momento dado como «literatura activa circulante» en el mismo. Por tanto, la obsolescencia se puede definir como la disminución con el tiempo de la utilización de la información, ya sea porque la información no se considera válida o porque, siendo válida, ha sido reemplazada por otra más moderna.

El semiperiodo consiste en la mediana de la distribución de las referencias por años de procedencia. El semiperiodo será tanto más corto cuanto más rápido sea el crecimiento de la ciencia en ese campo, ya que, si todos los artículos tienen la misma probabilidad de ser referenciados, aparecerá un mayor número de referencias a los trabajos más recientes, simplemente porque son más numerosos. Burton y Kebler^{3,37} comprobaron diferencias muy acusadas en las distintas disciplinas, que se debían a la existencia de dos tipos de literatura científica, la efímera y la clásica; cada una de ellas pesaba de modo diverso en las disciplinas. Hay materias con un fuerte componente de literatura clásica (matemáticas, geología, botánica), otras compuestas casi exclusivamente por literatura efímera

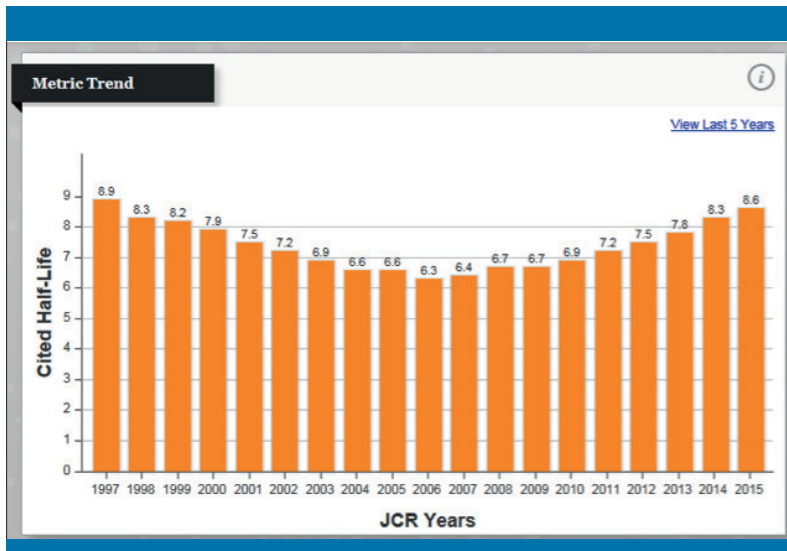


Figura 4. Evolución de la vida media de las citas en la revista *Pediatrics*. Fuente: *Journal Citation Reports*

(física, ingeniería) y algunas de carácter intermedio: química, fisiología o medicina clínica³.

Aunque la semivida no refleja el valor científico de una revista en particular, puede dar información sobre su política editorial o sobre el campo de la investigación: una semivida corta puede reflejar una política editorial que hace hincapié en la bibliografía más actual o en un campo de investigación en rápida evolución, mientras que una semivida larga puede reflejar que el énfasis se sitúa en la literatura más clásica. En la práctica, la semivida puede ser utilizada por los editores para ajustar su política editorial o para dar entrada a nuevas líneas de investigación¹⁶.

La vida media es proporcionada en el Journal Citation Reports de Web of Science. En la figura 4 se presenta el gráfico obtenido del Journal Citation Reports que muestra la vida media de la revista *Pediatrics* en los últimos 5 años.

Índice de Price

El índice de Price consiste en el porcentaje de referencias de menos de 5 años. El índice es muy alto en las ciencias puras, con componentes muy elevados de literatura efímera, y desciende a medida que aumenta la proporción de literatura clásica³.

Índice de aislamiento

Si se distribuyen las referencias por sus países de origen, puede obtenerse el índice de aislamiento, que es el porcentaje de referencias que corresponde a publicaciones del mismo país que las publicaciones citadoras. Las revistas médicas españolas tienen un bajo grado de aislamiento, ya que citan mayoritariamente revistas extranjeras. Algunos estudios previos³ sitúan este índice entre el 9 y el 14%. Otros países de nuestro entorno tienen un índice de aislamiento mayor, como Francia (17,5%), Alemania (17%), Gran Bretaña (37%) y Estados Unidos (70%), este último con un gran índice de aislamiento.

El bajo índice de aislamiento de las publicaciones españolas está muy influenciado por el elevado número de referencias a publicaciones angloamericanas y, en menor medida, a publicaciones de países europeos continentales. Por el contrario, las referencias a publicaciones asiáticas y de los países de América Latina son muy escasas.

Limitaciones de los indicadores basados en las citas

La utilización de indicadores basados en las citas en la evaluación de la actividad científica y la calidad de las publicaciones ha sido objeto de numerosas críticas y tiene varias limitaciones. Algunas de las más relevantes se describen a continuación^{2,3,38-42}:

1. Mientras que la citación de un trabajo demuestra su reconocimiento por la comunidad científica, la falta de citación no indica necesariamente la inutilidad del mismo, sobre todo porque para ser citado necesita como condición indispensable que esté «disponible y visible», es decir, que el trabajo haya sido difundido suficientemente, lo que no presenta una clara correlación con la calidad del mismo. Es bien conocido el hecho de que algunos investigadores muestran cierto pudor a citar trabajos aparecidos en revistas de países no desarrollados, aunque el trabajo les parezca de calidad.
2. Se ha demostrado que alrededor de un tercio de las citas son superficiales, es decir, no son citas fundamentales para el trabajo. En cambio, otros artículos relevantes sobre el tema no se citan. Esto último forma parte del llamado «fenómeno de obliteración», según el cual cuando un trabajo científico se hace tan genérico y está tan integrado en un campo que forma parte del cuerpo de conocimiento, no se cita explícitamente. Puede suceder en trabajos de alta calidad.
3. Los hábitos de citación varían mucho según los campos. Por ejemplo, los trabajos publicados en campos nuevos con mu-

cho potencial de crecimiento obtienen más citas que los trabajos de campos estáticos y reducidos.

4. Las citas no se pueden utilizar para comparar científicos de diferentes áreas, pues el tamaño del campo de investigación influye en el número de citas que un artículo es susceptible de recibir: en una comunidad de 50 investigadores, el número de citas, en valores absolutos, debe ser necesariamente más pequeño que en una comunidad de 5.000 investigadores.
5. Cada área científica tiene unos hábitos de citación diferentes. Mientras en bioquímica y en ciencias sociales se producen unas 30 referencias por artículo, en ingeniería sólo se generan 10 referencias y en matemáticas 5, por lo que la probabilidad de ser citado en literatura bioquímica es 6 veces mayor que en matemáticas.
5. Se producen desviaciones en la ética de la citación debido a que una considerable cantidad se omite por fallos de memoria, se plagia de otros trabajos sin haberlos leído, se hacen citas por sentido de lealtad a colegas próximos («yo te cito a ti si tú me citas a mí»), o existen presiones para citar determinados trabajos.

Conclusiones

Los indicadores basados en las citas se utilizan ampliamente en la evaluación de la actividad científica, a pesar de sus críticas y limitaciones, pues se basan en que los trabajos de cali-

TABLA 3

Decálogo para aumentar las citas de una revista

1. Proveer acceso por internet
2. Difundirla, a ser posible, en abierto o liberar los artículos pronto tras su publicación
3. Incluirla en el mayor número posible de bases de datos, sobre todo las que analizan las citas
4. Publicar artículos polémicos
5. Publicar revisiones
6. Publicar en inglés o bilingüe
7. Publicar sobre temas novedosos y de actualidad
8. Invitar a que los autores muy citados publiquen en la revista
9. Enviar noticias a los medios sobre temas de interés publicados en la revista
10. Recomendar a los autores que citen trabajos publicados en la misma revista

TABLA 4

Veinte estrategias para aumentar las citas de los investigadores

1. Citar trabajos anteriores cuando sean relevantes para un nuevo manuscrito, pero no hacerlo por el simple hecho de querer aumentar la cuenta de citas
2. Elegir cuidadosamente las palabras clave en los manuscritos para que éstos aparezcan en las búsquedas en bases de datos. Usar exactamente las palabras clave por las que buscarían los investigadores en el campo
3. Utilizar estas palabras clave en el título y repetidamente en su resumen. La repetición de palabras clave aumentará la probabilidad de que el manuscrito esté en la parte superior de las listas de trabajos que proporcionan los motores de búsqueda, por lo que será más probable que se lea nuestro trabajo.
4. Utilizar una firma uniforme en todos los trabajos. Esto facilitará que otros encuentren todos nuestros trabajos publicados. Considerar obtener un identificador de investigación, como el ORCID, el ResearcherID de WoS o el de Scopus. Este identificador se puede proporcionar en la firma de correo electrónico y vincularlo a nuestra lista de publicaciones para que cualquier persona tenga acceso a nuestras publicaciones
5. Utilizar una filiación institucional uniforme en todos los trabajos, evitando las abreviaturas y siglas. Es esencial proporcionar detalles exactos de contacto (dirección para la correspondencia y dirección de correo electrónico) para que los investigadores puedan contactar directamente para cualquier consulta, información adicional y discusiones sobre el trabajo publicado
6. Comprobar que el nombre y la filiación son correctos en las pruebas finales del manuscrito y que la información del documento es precisa en las búsquedas en la base de datos
7. Publicar en revistas indexadas en muchas bases de datos
8. Publicar en revistas con un alto factor de impacto, pues se ha demostrado que los artículos publicados en ellas incrementan la probabilidad de ser citados
9. Conocer las revistas en las que se quiere publicar y preguntarse si los investigadores en el campo publican también allí
10. Hacer que el manuscrito sea fácilmente accesible. Si el trabajo no se publica en acceso abierto, publicar los *preprints* o alguna versión previa definitiva en un repositorio
11. Compartir los datos, ya que existe evidencia de que compartir los datos puede aumentar las citas. Para ello, cabría considerar publicarlos en sitios web de intercambio de datos, como Dryad, Figshare o SlideShare. Como mínimo, autoarchivarlo en nuestra web personal
12. Presentar el trabajo en congresos. Aunque las presentaciones en congresos no son citadas por otros, permitirán que su investigación sea más visible para las comunidades académicas y de investigación
13. Utilizar las redes sociales proporcionando enlaces a los artículos en Facebook, Twitter, Academia.edu, ResearchGate, Mendeley y también en nuestra página web curricular
14. Promover activamente nuestro trabajo, hablando de él con otros investigadores y enviando copias por correo electrónico a los que puedan estar interesados
15. También puede ser útil crear un blog o un sitio web dedicado a nuestras investigaciones y compartirlo. Mantenerlo siempre actualizado
16. Publicar con autores internacionales. El análisis de citas demuestra que los documentos con coautores internacionales se citan hasta 4 veces más a menudo
17. Escribir algún artículo de revisión exhaustivo y sustancial, pues recibe más citas que los artículos originales
18. Publicar tutoriales, es decir, documentos que organizan e introducen el trabajo en el campo. Un tutorial asume que su audiencia es inexperta, por lo que resalta los conceptos básicos del campo y proporciona ejemplos concretos que encarnan estos conceptos
19. Intentar publicar una noticia en los medios de comunicación
20. Elaborar una versión divulgativa del manuscrito que permita publicarlo en medios no científicos

dad son citados, mientras que los que no son relevantes se ignoran. Una de sus limitaciones más importantes es la falta de exhaustividad de las fuentes que analizan las citas, con un marcado sesgo a favor de las publicaciones anglosajonas, lo que perjudica la visibilidad y la citación de las revistas de numerosos países, como España⁴³⁻⁴⁵. Por otra parte, el uso sistemático de estos indicadores ha fomentado la picaresca en torno a la citación, y que tanto las revistas como los investigadores adopten estrategias para aumentar sus citas. En la tabla 3 se presenta un decálogo para aumentar las citas de las revistas y en la tabla 4 se enumeran 20 estrategias para incrementar las citas de los investigadores. ■

Bibliografía

- Aleixandre-Benavent R, González de Dios J, Castelló Cogollos L, Navarro Molina C, Alonso-Arroyo A, Vidal-Infer A, et al. Bibliometría e indicadores de actividad científica (1). La evaluación de la investigación y de la actividad científica en pediatría a través de la bibliometría. *Acta Pediatr Esp.* 2017; 75: 18-25.
- González de Dios J, Aleixandre Benavent R. Evaluación de la investigación en biomedicina y ciencias de la salud: indicadores bibliométricos y cibernéricos. *Bol Pediatr.* 2007; 47: 92-110.
- López Piñero JM, Terrada ML. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (III). Los indicadores de producción, circulación y dispersión, consumo de la información y repercusión. *Med Clin (Barc).* 1992; 98: 141-148.
- Bolaños-Pizarro M, Navarro-Molina C, Alonso-Arroyo A, González de Dios J, Aleixandre-Benavent R. Fuentes de información bibliográfica (IV). Bases de datos bibliográficas multidisciplinarias de interés en pediatría: Science Citation Index y Journal Citations Reports (Web of Science) (parte 1). *Acta Pediatr Esp.* 2011; 69: 291-298.
- Bolaños-Pizarro M, Navarro-Molina C, Alonso-Arroyo A, González de Dios J, Aleixandre-Benavent R. Fuentes de información bibliográfica (IV). Bases de datos bibliográficas multidisciplinarias de interés en pediatría: Science Citation Index y Journal Citations Reports (Web of Science) (parte 2). *Acta Pediatr Esp.* 2011; 69: 343-352.
- Navarro-Molina C, González-Alcaide G, Bolaños-Pizarro M, González de Dios J, Aleixandre-Benavent R. Fuentes de información bibliográfica (VI). Obtención de literatura científica con la base de datos Scopus y los buscadores especializados Scirus y Google Académico. *Acta Pediatr Esp.* 2011; 69: 131-136.
- Aleixandre-Benavent R. Bibliometría e indicadores de actividad científica. En: Jiménez Villa J, Argimón Pallás JM, Martín Zurro A, Vilardell Tarrés M, eds. *Publicación científica biomédica: cómo escribir y publicar un artículo de investigación.* Barcelona: Elsevier España, 2010; 363-384.
- Bordons M, Zulueta MA. Evaluación de la actividad científica a través de indicadores bibliométricos. *Rev Esp Cardiol.* 1999; 52: 790-800.
- Campanario JM. Empirical study of journal impact factors obtained using the classical 2-year citation window versus a five-year citation window. *Scientometrics.* 2011; 87: 189-204.
- Adams J. Early citation counts correlate with accumulated impact. *Scientometrics.* 2005; 63: 567-581.
- Dorta-González P, Dorta-González MI. Comparing journals from different fields of science and social science through a JCR subject categories normalized impact factor. *Scientometrics.* 2003; 95: 645-672.
- Vanclay JK. Ranking forestry journals using the h-index. *J Informetr.* 2008; 2: 326-334.
- Waltman NJ, Van Eck TN, Van Leeuwen MS, Visser AFJ, Van Raan. Towards a new crown indicator: an empirical analysis. *Scientometrics.* 2011; 87: 467-481.
- Jacso P. Five-year impact factor data in the Journal Citation Reports. *Online Inf Rev.* 2009; 33: 603-614.
- Waltman L. A review of the literature on citation impact indicators. *J Informetr.* 2016; 10: 365-391.
- Durieux V, Gevenois PA. Bibliometric indicators: quality measurements of scientific publication. *Radiology.* 2010; 255: 342-351.
- Albarrán P, Ortuño I, Ruiz-Castillo J. High-and low-impact citation measures: empirical applications. *J Informetr.* 2011; 5: 122-145.
- Seglen PO. The skewness of science. *J Am Soc Inf Sci.* 1992; 43: 628-638.
- Bornmann L. How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Res Eval.* 2014; 23: 166-173.
- Tijssen RJ, Visser MS, Van Leeuwen TN. Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics.* 2002; 54: 381-397.
- Hirsch JE. An index to quantify an individual's scientific research output. *Proc Nat Acad Sci.* 2005; 102: 16.572.
- Aznar J, Guerrero E. Análisis del índice h y propuesta de un nuevo índice bibliométrico: el índice global. *Rev Clin Esp.* 2010; 211: 251-256.
- Costas R, Bordons M. Una visión crítica del índice h: algunas consideraciones derivadas de su aplicación práctica. *Prof Inf.* 2007; 427-432.
- Van Raan AFJ. Comparisons of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics.* 2006; 67: 491-502.
- Grupo Scimago. El índice h de Hirsch: aportaciones a un debate. *Prof Inf.* 2006; 15: 304-306.
- Kelly CD, Jennions MD. The h-index and career assessment by numbers. *Trends Ecol Evol.* 2006; 21: 167-170.
- Glänzel W. On the h-index: a mathematical approach to a new measure of publication activity and citation impact. *Scientometrics.* 2006; 67: 315-321.
- Jin B, Liang L, Rousseau R, Egghe L. The R- and AR-indices: complementing the h-index. *Chinese Sci Bull.* 2007; 52: 855.
- Egghe L. Theory and practice of the g-index. *Scientometrics.* 2006; 69: 131-152.
- Egghe L. Mathematical theory of the h-and g-index in case of fractional counting of authorship. *J Am Soc Inf Sci Technol.* 2008; 59: 1.608-1.616.
- Jin BH. The AR-index: complementing the h-index. *ISSI Newsletter.* 2007; 3: 6.
- Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F. h-Index: a review focused in its variants, computation and standardization for different scientific fields. *J Informetr.* 2009; 3: 273-289.
- Egghe L. The Hirsch index and related impact measures. *Ann Rev Inf Sci Technol.* 2010; 44: 65-114.
- Norris M, Oppenheim C. The h-index: a broad review of a new bibliometric indicator. *J Doc.* 2010; 66: 681-705.
- Lundberg J. Lifting the crown-citation z-score. *J Informetr.* 2007; 1: 145-154.
- Moed HF, De Bruin RE, Van Leeuw TN. New bibliometric tools for the assessment of national research performance: database des-

- cription, overview of indicators and first applications. *Scientometrics*. 1995; 33: 381-422.
37. Burton RE, Kebler RW. The "half-life" of some scientific and technical literatures. *Am Doc*. 1960; 11: 18-22.
 38. González de Dios J, Sempere AP, Aleixandre-Benavent R. Las publicaciones biomédicas en España a debate (II): Las «revoluciones» pendientes y su aplicación a las revistas neurológicas. *Rev Neurol*. 2007; 44: 101-112.
 39. Aleixandre Benavent R, Valderrama Zurián JC, Castellano Gómez M, Simó Meléndez R, Navarro Molina C. «Archivos de Bronconeumología»: una de las 3 revistas médicas españolas con mayor factor de impacto nacional. *Arch Bronconeumol*. 2004; 40: 563-569.
 40. González Alcaide G, Castellano Gómez M, Valderrama Zurián JC, Aleixandre Benavent R. Literatura científica de autores españoles sobre análisis de citas y factor de impacto en biomedicina (1981-2005). *Rev Esp Doc Cient*. 2008; 31: 344-365.
 41. Aleixandre R, Giménez Sánchez JV, Terrada Ferrandis ML, López Piñero JM. Análisis del consumo de información en la revista «Atención Primaria». *Aten Primaria*. 1996; 17: 321-325.
 42. Aleixandre Benavent R, Porcel Torrens A. El factor de impacto de las revistas científicas. *Trast Adict*. 2000; 1: 264-271.
 43. González Alcaide G, Valderrama Zurián JC, Aleixandre-Benavent R. Análisis del proceso de internacionalización de la investigación científica española. *Rev Esp Doc Cient*. 2012; 35: 94-118.
 44. Aleixandre R, Valderrama JC, Alonso-Arroyo A, Miguel-Dasit A, González de Dios J, De Granda Orive JI. Español vs inglés como idioma de publicación de neurología. *Neurología*. 2007; 22: 19-26.
 45. Alonso Arroyo A, González de Dios J, Bolaños Pizarro M, Castelló Cogollos L, González Alcaide G, Navarro Molina C, et al. Análisis de la productividad e impacto científico de la pediatría española (2006-2010). *An Pediatr (Barc)*. 2013; 78(6): 409.e1-409.e17.