

## El índice-H y Google Académico: una simbiosis cienciométrica inclusiva

### H-index and Google Scholar: a inclusive scientometrics symbiosis

Dr. Luis Carlos Silva Ayçaguer

Centro Nacional de Información de Ciencias Médicas-Infomed. La Habana, Cuba.

---

#### RESUMEN

Se examina el origen del "factor de impacto" como indicador de la influencia de una revista científica y su evolución hasta convertirse en un instrumento altamente redituable para la empresa *Thomson Reuters*, gracias en parte a su naturaleza excluyente y anglocéntrica. El índice-H de Hirsh y sus variantes, así como la posibilidad de conjugarlo con las prestaciones de Google Académico, son valorados con detalle y conceptuados como componentes de un modelo alternativo, con imperfecciones diversas, pero libre de buena parte de las objeciones que se han hecho a los indicadores precedentes. Tales posibilidades resultan especialmente atractivas cuando se involucran idiomas diferentes al inglés en el análisis.

**Palabras clave:** factor de impacto, índice-H, *Hirsh*, *Thomson Reuters*, cienciometría, Google Académico.

---

#### ABSTRACT

The paper shows how the dominant journal impact factor has arrived to its present features and discuss the degree in which this metric is prone to be manipulated and misused, as opposed to the prescribed utilization by *Thomson Reuter*, the corporation in charge of its official computation. *Hirsch's H-index* and a large family of related indicators seek to give a single number that in some sense summarizes an author's research output and its impact. The free public availability of information offered by Google Scholar allows citation counts, and analyses based thereon, to be performed and duplicated by anyone. Combining H-index with this information provides an avenue for more transparency and supply an extraordinary

---

opportunity to develop a fairer scientometric analysis, specially when languages other than English are involved.

**Key words:** impact factor, H-index, *Hirsch*, *Thomson Reuters*, scientometrics,, Google Scholar.

---

## ANTECEDENTES Y EL FACTOR DE IMPACTO

La producción científica es un componente medular de la cultura de la sociedad y un elemento esencial en el proceso de generación y evaluación de tecnologías. Tales niveles de protagonismo en la sociedad hacen que resulte imperativo contar con recursos que faciliten la valoración objetiva de sus méritos. La globalización, por su parte, impone que los indicadores empleados con ese fin tengan alcance operativo universal. El impacto y la validez de dicha producción se mide para establecer jerarquías de prestigio científico y, por esa vía, pautas orientativas para autores, editores y otros actores sociales entre los que se destacan gestores e inversores en materia de salud, así como productores de medicamentos o vacunas, equipos y servicios sanitarios. Tal realidad permite comprender fácilmente que las implicaciones propias de un indicador dado desbordan lo meramente operacional y epistemológico; su teleología alcanza una connotación económica, social y, finalmente, política. El análisis de las alternativas que se han sugerido para tal medición, además del examen de sus valores métricos, exige de un enfoque sociocultural que reconozca el carácter histórico y contextual de las propuestas.

Entre los indicadores bibliométricos sugeridos para cuantificar la influencia de una revista científica, el más afamado es, sin dudas, el popularmente conocido como *factor de impacto* (FI). Fue concebido hace más de medio siglo por los documentalistas norteamericanos *Eugene Garfield* e *Irving Sher*<sup>1</sup> con el nombre formal de *journal impact factor* (JIF). En términos muy generales puede describirse como la razón entre el número medio de citas de que fueron objeto durante cierto período los artículos publicados en la revista y el número de artículos publicados por dicha revista a lo largo de un lapso previo al período en que se produjeron tales publicaciones.

El FI, sin embargo, ha tenido su propia evolución. De la propuesta original a la actualidad, ha experimentado importantes modificaciones, no tanto en la estructura de su formulación como en la forma de determinar sus componentes.

Originalmente, el FI no establecía un marco temporal fijo. Fue en 1968 que dos especialistas británicos introdujeron una acotación<sup>2</sup> que poco después fuera "bendecida" por el propio *Garfield* y su empresa.<sup>3</sup> Tal restricción, que retrotrae la contabilización de artículos a solo dos años previos, distaba de ser "inocente". Como demuestran los investigadores canadienses *Eric Archambault* y *Vincent Lariviere* con argumentos que no reproduciré aquí por razones de espacio pero que pueden ser consultados en un trabajo reciente,<sup>4</sup> ese lapso fue calculadamente elegido. Con él, las publicaciones producidas en inglés serían beneficiarias de un mayor factor de impacto, contribuyendo así a un bucle que se arrastra hasta la actualidad: puesto que las revistas anglófonas eran las que ostentaban mayor FI, resultaba más redituable para los investigadores publicar en ellas, lo cual a su vez

garantizaba a dichas revistas un mayor FI en el futuro. A esta maniobra se sumaron otras, bastante más escabrosas, como se muestra a continuación.

Cuando se calcula el factor de impacto  $FI_N$  para el año N en una revista específica, tal y como se hace en la actualidad, se procede del modo siguiente. Se calculan los valores  $P(TR)_{N-1}$  y  $P(TR)_{N-2}$ , que representan la cantidad de artículos "citables" que hayan sido publicados en esa revista a lo largo de los años N-2 y N-1 respectivamente, así como el número de citas que a lo largo del año N recibieron todos los artículos publicados en dicha revista. Si llamamos  $C(TR)_N$  a ese último número, el factor de impacto es la razón que se representa en la fórmula siguiente:

$$FI_N = \frac{C(TR)_N}{P(TR)_{N-2} + P(TR)_{N-1}}$$

La comprensión de la fórmula exige algunas importantes precisiones. Las letras TR que figuran entre paréntesis se han incluido en cada uno de los tres términos involucrados para destacar que no cualquier cita recibida es válida para ser contabilizada en el numerador, ni cualquier artículo publicado puede hacer aportes a los sumandos del denominador. Las citas del numerador han de proceder de revistas indexadas por la compañía *Thomson Reuters Scientific Inc.* y los artículos del denominador solo son aquellos que la susodicha compañía considera "citables" entre los publicados a lo largo de los dos años anteriores al momento del cómputo.

Para fijar ideas, supongamos, por ejemplo, que en el año 2004 una revista ha publicado 43 artículos que *Thomson Reuters* (TR en lo sucesivo) considera "citables", y que ese número ascendió a 38 en el 2005. Si las revistas indexadas por esta empresa han realizado durante 2006 un acumulado de 237 citas a trabajos publicados por la revista en cuestión, entonces su factor de impacto para el 2006 sería:

$$FI_{2006} = \frac{C(TR)_{2006}}{P(TR)_{2004} + P(TR)_{2005}} = \frac{237}{43+38} = 2,926$$

No pocas críticas ha recibido el FI desde su creación.<sup>5-7</sup> En principio, que se produzcan objeciones es algo natural: cualquier propuesta innovadora las recibe como parte del "toma y daca" que le es consustancial al desarrollo científico. Todo procedimiento seminal es objeto de propuestas modificadoras o complementarias para mejorar su desempeño, que surgen tanto de la reflexión teórica como de las experiencias que van acopiando sus usuarios.

En su aplicación contemporánea, sin embargo, el FI despierta suspicacias especiales. La más comprometedor y delicada concierne al hecho de que su fórmula esté tan medularmente atravesada por decisiones que adopta *Thomson Reuters*. Esta mega-empresa de productos informativos, bibliométricos en particular, funciona con la lógica del mercado, como cualquier transnacional. Históricamente ha ido tragándose a sus competidores desde que en 1992 adquirió al *Institute for Scientific Information (ISI)* que dirigiera el propio *Eugene Garfield* y más recientemente a la Agencia Reuters (de ahí que haya pasado de denominarse *ScientificThomson* a ostentar su nombre actual) y obtiene enormes beneficios que, por poner un ejemplo, ascendieron solo en 2002 a 7600 millones de dólares.<sup>8</sup>

El problema fundamental reside no solo en que el indicador es rehén de las normas establecidas por ella sino que tales normas son en buena medida opacas. La falta de transparencia más importante estriba en que el concepto de "artículo citable" carece de una definición pública e inequívoca, y la empresa se niega a comunicarla.

De hecho, si alguien desea computar el FI "entendido en su versión oficial" simplemente le será imposible. Solo puede remitirse a admitir el que la empresa *Thomson Reuters* publica regularmente.

Por otra parte, el FI tiene un rasgo adicional que favorece la aplicación de maniobras que adulteran el sentido primigenio con que fue creado. Hay una asimetría notoria entre el tipo de documentos que se contabilizan en el numerador (referencias a cualquier tipo de material) y en el denominador de la ecuación (solo los susodichos "citables"), la cual propicia y consiente la "reingeniería" de las revistas sobre el tipo de documentos que publican: basta incrementar la aparición de artículos que se consideren "no citables" (o conseguir que TR reduzca ese número) para aumentar artificialmente el FI. Si, por ejemplo, como resultado de estas estrategias se consiguiera que el número de artículos citables se redujera en 5 para cada uno de los años del ejemplo que se ha puesto, el FI de esa revista hipotética pasaría de ser 2,926 a 3,338.

Pudiera pensarse que tal artificio lo pueden consumir todas las revistas, pero ocurre que el carácter de citable o no es materia de posibles negociaciones, algo que solo las revistas más poderosas estarían en condiciones de llevar adelante.

No son detalles menores, pues confieren una elasticidad al FI que se presta a manipulaciones de las que han sido acusadas tanto algunas revistas como la propia empresa.<sup>9,10</sup> En un editorial de la revista *Clinics* se ponen varios ejemplos que, según el autor, demuestran la creciente presión a que están sometidas las revistas para recomponer el factor de impacto, favorecida por la naturaleza oculta de las definiciones de sus términos y por la disposición de *Thompson Reuters* a negociar con los editores.<sup>11</sup>

Así, lo que empezó por ser un intermediario metodológico para medir indirectamente la calidad de una revista, ha terminado por convertirse en una meta que distorsiona las decisiones en materia de publicación, hoy más en manos de administradores de la ciencia que de los propios científicos.

Es obvio que las características de un indicador y, sobre todo, las del modo en que se emplea, pueden ayudar a conocer la realidad; pero su naturaleza, y sobre todo las reglas que rigen para aplicarlo, pueden a su vez ir modificando la realidad que se pretende caracterizar con él.

El descontento con este estado de cosas se ha expresado con fuerza en años recientes.<sup>12</sup> Sobre estas bases, los editores de la revista *PLoS Medicine* han llegado a concluir que "la ciencia está en la actualidad valorada por un proceso que es en sí mismo subjetivo, no científico y subrepticio".<sup>13</sup> Algunas voces van más lejos, como por ejemplo, la que desde la revista *Epidemiology* afirmaba que "Los científicos deberían estar muy irritados de que el valor de la ciencia sea medido mediante una métrica secreta...[y] deberían renunciar al factor de impacto de *Thomson Reuters* así como procurarse una métrica universal que siga los principios de la ciencia: precisión, reproducibilidad, y naturaleza pública y accesible a los datos y los análisis".<sup>14</sup>

Cabe consignar que, pese a los muchos juicios críticos que ha merecido, el FI es defendido con vehemencia por la citada empresa e incluso por su octogenario

creador.<sup>15,16</sup> Lo cierto es que el *Thomson ISI Web of Knowledge*, sitio operativo a estos efectos de la empresa arriba caracterizada, y en particular sus instrumentos *Science Citation Index* y *Journal Citation Reports* proveen cada año los susodichos *impact factors*. El FI, en suma, es aún ampliamente dominante como recurso valorativo, de modo que no se vislumbra, ciertamente, su erradicación del panorama cuantitativo contemporáneo. *Archambault* y *Lariviere* en el trabajo arriba citado escriben textualmente:<sup>4</sup>

"No podemos ser ingenuos en torno a las reglas de *Thomson Scientific*. Aunque el mensaje oficial es que el JCR no debe emplearse para evaluar el resultado de las investigaciones, permitiendo que por su conducto se concedan promociones académicas, todo el mundo sabe que esta es la situación prevaleciente y que *Thomson Scientific* se beneficia tanto económicamente como en términos de prestigio por el extendido uso que se hace de JCR. Incluso, conociendo las insistentes críticas de que es objeto este indicador, la empresa no lo ha modificado con vistas a minimizar sus sesgos y su vulnerabilidad respecto de las manipulaciones".

Más allá de todas estas consideraciones, el mero sentido común conduce a sospechar que el FI, a casi seis décadas de su invención y 45 años después de su puesta en práctica en su forma actual, no puede mantenerse como un recurso adecuado para medir el impacto de la información científica, especialmente si se repara en que a lo largo de los decenios transcurridos se han verificado reemplazos tecnológicos y conceptuales trascendentales en materia de acceso a la información científica y de las modalidades que ésta asume. Basta reparar en la vigencia del movimiento de acceso abierto y en la existencia de repositorios institucionales, entre otras realidades impensables hace solo unos años.<sup>17</sup>

Consecuentemente con todo lo anterior, en los últimos años se ha producido un nutrido conjunto de propuestas alternativas o complementarias. Una de ellas, el llamado "índice-H" ha resultado ser la más exitosa y atractiva de tales manifestaciones. En las próximas secciones procedo a exponer y documentar algunos de sus rasgos fundamentales y su impacto en el mundo cuantitativo.

## EL ÍNDICE-H

El nombre con que se ha popularizado alude a su creador, *Jorge E. Hirsch*, físico de la Universidad de California, San Diego, EE. UU. Sus enérgicas advertencias antibelicistas motivadas por lo que consideraba en 2005 como "agresivas e innecesarias políticas militares de EE. UU" le dieron renombre y cierta aureola de intelectual contestatario. Hirsch había cobrado notoriedad en su disciplina, sin embargo, desde que en 1989 presentara en un congreso un trascendente trabajo. En él se refutaba la teoría por entonces generalmente aceptada de los superconductores de baja temperatura (la llamada teoría BCS). Si bien era muy reconocido por sus colegas, conocía que "debía" publicar su excelente resultado en revistas con alto FI. Sin embargo, la lógica implícita en su mecanismo de aplicación no era compatible con su convicción de que la clave no estaba tanto en dónde se publicaban sus trabajos como en el grado en que ellos eran citados por la comunidad científica. Según relata el analista *Guy Gugliotta*,<sup>18</sup> destinó entonces dos años a examinar el problema y a concebir una alternativa. De este marco teórico y epistemológico, surgió su ingenioso indicador, propuesto en el propio año 2005,<sup>19</sup> al cual podemos considerar también un indicador contestatario.

No exento de limitaciones, que más adelante se esbozan, el índice-h constituye un notable aporte orientado a medir la productividad de un científico. Por su naturaleza, es igualmente aplicable a una revista completa, e incluso a una institución.<sup>20</sup>

Desde el punto de vista métrico, su mérito fundamental reside en que intenta aunar la cantidad de la producción con su calidad. El problema básico que aqueja a los indicadores bibliométricos tradicionales, tales como el número total de artículos o el número de citas recibidas, estriba en que la primera de tales medidas no repara en la calidad de los trabajos publicados —indirectamente medida por el número de citas—, mientras que la segunda puede ser alta para individuos que han hecho muy pocas publicaciones, siempre que estas hayan sido muy citadas. El índice-H procura condensar en una única magnitud el número de publicaciones y el de citas que ellas reciben.

Para introducir su definición, consideremos el caso de un investigador que ha publicado  $N_p$  artículos. El valor del índice es un número entero que asciende a  $t$  si  $t$  de sus  $N_p$  trabajos publicados ha recibido al menos  $t$  citas cada uno, a la vez que los otros  $(N_p - t)$  trabajos han sido citados como máximo en  $t$  ocasiones cada uno.

Algunos ejemplos pueden ser útiles para comprender la definición y para aquilatar de manera más acabada las virtudes y los defectos de este indicador. La [tabla](#) recoge varias situaciones hipotéticas.

**Tabla 1.** Producción científica, citas recibidas e índice-h para un investigador a lo largo de seis años consecutivos

Artículos	Citas 1995	Citas 1996	Citas 1997	Citas 1998	Citas 1999	Citas 2000
1	5	12	24	24	29	29
2	5	11	12	12	16	16
3	4	10	10	10	15	15
4	3	8	8	8	13	13
5	2	6	6	6	11	11
6	1	6	6	6	11	11
7	1	5	5	5	5	10
8	1	3	3	3	3	8
9	-	2	2	2	2	7
10	-	-	2	2	2	7
11	-	-	-	1	1	6
Índice-h	3	6	6	6	6	8

El lector puede descargar un sencillo programa confeccionado en Excel por el autor, con el cual hacer sus propios cálculos y sondeos con vistas a entender mejor cómo opera el indicador. Con él puede corroborarse que en el año 1998, este investigador tenía publicados 11 trabajos y el valor correspondiente del índice era igual a 6; no se modificó al año siguiente a pesar de que cada uno de los 6 trabajos más citados recibió 5 citas adicionales. En cambio, cuando ese mismo incremento

se produjo en el 2000 para todos sus trabajos respecto de 1998, el índice-H ascendió a 8. En 1995, había sido igual a 3, y subió a 6 al año siguiente como resultado del incremento en las citas de sus 8 trabajos iniciales; sin embargo, durante los tres años siguientes el índice se mantuvo en 6 a pesar de que año tras año se fue incrementando tanto su producción como el impacto que esta iba teniendo, hasta que en el año 2000 tal incremento alcanzó una magnitud sustantiva.

Desde el punto de vista práctico, y en términos reales, su mayor atractivo reposa en que cualquiera puede calcularlo de manera independiente, tanto para un autor como para una revista. El listado de publicaciones y el número de citas no son difíciles de obtener en principio: basta poner el nombre del autor o de la revista en el buscador de *Google Académico* y el buscador devuelve esta información. Sin embargo, el acopio ha de ser en principio algo artesanal, navegando página por página si hay muchas entradas. Por otra parte, comparece el problema de que las obras de un mismo autor pueden tener la autoría marcada según diferentes denominaciones (Juan Pérez, Juan José Pérez, Juan Pérez García, etc.), un problema acerca del cual se ha alertado cuando se insiste en que cada autor "normalice" su firma, a la vez que pueden existir artículos de diferentes personas con el mismo nombre. También las revistas pueden figurar con diferentes nombres (por ejemplo, con el nombre completo o solo por su sigla). No obstante, aunque laborioso, es posible hacerlo. No dependemos de información ubicada en una caja negra a la que no se tiene acceso: basta acudir a recursos al alcance de todos, hasta hace muy pocos años inexistentes.

### ALGUNAS CARACTERÍSTICAS DEL ÍNDICE-H

A continuación se enumeran algunos rasgos peculiares del indicador, parte de los cuales emergen claramente de los ejemplos incluidos en la tabla.

- Es imposible que un investigador o una revista tengan un alto valor en el índice-H si no tiene un número importante de artículos publicados. De hecho, ese número es el máximo valor que pudiera alcanzar el índice. Pero no basta con tener muchos artículos publicados para que este sea elevado. Puede ocurrir que  $N_p$  sea igual a 80, pongamos por caso, y que sin embargo el índice-h asociado sea 5, un valor que pudiera alcanzarse con solo 5 artículos publicados, siempre que todos ellos hayan sido citados cinco o más veces.
- A efectos de comparación entre investigadores (o revistas), el índice pierde eficacia si no proceden de la misma disciplina, pues los patrones de citación de los trabajos suelen diferir de un campo a otro.
- En principio, el índice-H puede calcularse sin contemplar restricción temporal alguna. Es decir, se puede aplicar a toda la obra publicada por un autor o por una revista. En tal caso, tiene una singularidad: el valor puede incrementarse con el tiempo, pero jamás podrá reducirse. Sin embargo, también puede ceñirse el cómputo a un período específico (por ejemplo, a los últimos 5, 10 o 15 años). Con esto, se estaría cuantificando la productividad *actual* del ente valorado, y esta sí pudiera decrecer con el tiempo.
- Hay consenso, y así lo señalaba el propio *Hirsh* en su artículo seminal, en que el índice-H no es apropiado para medir el desempeño académico de profesionales muy jóvenes, quienes no han tenido tiempo ni de alcanzar una producción cuantitativamente considerable, ni de acumular citas que verosímelmente se producirán en años venideros.

- Los artículos de revisión suelen recibir mayor cantidad de citas que los originales; consecuentemente, una persona o revista que predominantemente publicara revisiones obtendría un índice-H mayor que el de aquellas que publican fundamentalmente trabajos originales.
- El índice-H desdeña la importancia de trabajos muy singulares. Dos personas o revistas pueden tener el mismo valor "por ejemplo, igual a 27" a pesar de que uno de ellas publicó, digamos, un trabajo que fue objeto de 340 citas, mientras que ninguno de los trabajos de la otra ha recibido más de 27 citas.

Algunos de estos rasgos han sido conceptuados como deficiencias que le aquejan. Diversas adaptaciones y variantes se han creado para solventar las limitaciones que se han atribuido al indicador, tal y como se esboza en la siguiente sección.

### ALTERNATIVAS BASADAS EN EL ÍNDICE-H

La avenida que *Hirsh* abrió para la investigación y la innovación ha sido extraordinariamente fértil. La primera alternativa fue introducida en su trabajo original por el propio *Hirsh* para encarar el problema de que los investigadores jóvenes estarían en "desventaja", ya que el valor de su índice-H estaría acotado por su aún escasa producción, así como por el hecho de que sus contribuciones estarían en proceso creciente de citaciones. Surge así el *m-quotient*, resultado de dividir el índice-H por el número de años que lleva el autor en su producción, medido a través del lapso transcurrido desde su primera publicación.

Entre las variantes más exitosas, se halla la que fuera propuesta por el profesor *Leo Egghe*<sup>21</sup> y bautizada como índice-g (*g-index*). Si bien los artículos más reconocidos por otros autores tienen gravitación sobre el índice-H, una vez que figuran entre los t más citados, el número de citaciones es irrelevante. La propuesta de *Egghe* confiere, en cambio, más peso a los artículos más recurrentemente citados, mediante un recurso algo complicado, pero que ha resultado ser altamente efectivo\*.

*Rousseau*<sup>22</sup> introdujo el concepto de "núcleo-*Hirsh*" (*h-core* en inglés), con el que se alude al conjunto de los t artículos más citados de esta revista, donde t es el valor al que asciende el índice-H. En lo que sigue, llamaremos núcleo-H a dicho conjunto. Sobre esta base, *Bihou Jin*, editor principal de la revista *Science Focus*, realizó otra atractiva propuesta consistente en promediar las citas recibidas por los artículos que se hallan en el Núcleo-H<sup>23</sup> y conocida como el *a-index*, donde la "a" alude, precisamente, a que se trata de un promedio (*average*).

He mencionado solo un par de variantes relevantes, pero realmente se ha producido una verdadera cascada de propuestas inspiradas en el trabajo original de *Hirsh*.<sup>24-37</sup> A continuación nos detendremos en la modalidad introducida y puesta en práctica por Google Académico.

### GOOGLE ACADÉMICO Y EL ÍNDICE-H

*Google Académico* (en inglés, *Google Scholar*, GS) es un buscador de Google, lanzado al público en versión Beta en noviembre de 2004, circunscrito al ámbito académico, y soportado por una base de datos libremente disponible en Internet.

---



Allí se almacena un extenso conjunto de trabajos de investigación científica incluyendo el universo de acceso abierto y procedente de las más diversas disciplinas.

Su rasgo distintivo más notable es que, aunque se especializa en artículos científicos e incluye las revistas más connotadas en el mundo, también abarca trabajos en diversos formatos de presentación (libros, documentos presentados en conferencias científicas, tesis e informes técnicos ubicados en repositorios institucionales de universidades y centros de investigación) y, en particular, en revistas no indexadas por *Thomson Reuters*.

La trascendencia de esta singularidad es enorme, especialmente para investigadores cuyos aportes, "al menos en parte", ven la luz en publicaciones nacionales o regionales, con frecuencia en idiomas diferentes del inglés. Deliberadamente o no, lo cierto es que GS se ha convertido en una alternativa no excluyente y, en esa medida, antihegemónica respecto de los polos de poder prevalecientes hasta ahora.

Esto ha llevado incluso a que una de las especialistas más destacadas en materia bibliométrica, la Dra. *Anne Will Harzing*, profesora de la Universidad de Melbourne, expresara que GS abre las puertas a la "democratización" del análisis de citas<sup>38</sup> bajo el argumento adicional de que los rasgos de GS se complementan excelentemente con las características del índice-H.

Una pregunta natural en este contexto es la siguiente: ¿qué se pierde y qué se gana cuando se emplea el índice asociado con la base disponible en GS respecto del uso de fuentes convencionales tales como *Scopus* o la *Web of Science* y su factor de impacto? Lo primero que cabe decir es que estas fuentes más convencionales o históricas no tienen por qué conceptuarse como estándares para la valoración. Que GS dé lugar a resultados diferentes es lógico y esperable, pues la fuente informativa de partida es muy distinta, quizás con elementos polémicos, pero sin duda más completa. El solo hecho de que GS contemple otros idiomas en pie de igualdad al inglés, así como otras fuentes documentales, permite vaticinar estas diferencias y a la vez constituye un rasgo encomiable en la medida que atempera el sesgo que supone la discriminativa hegemonía idiomática del inglés. Consecuentemente, algunos estudios dan cuenta de diferencias considerables.<sup>39,40</sup> Sin embargo, lo atribuyen a limitaciones de GS, algo sumamente discutible por la razón apuntada. En este contexto se ha llamado la atención<sup>41</sup> acerca de revistas sumamente prestigiosas en sus respectivos ámbitos disciplinares que se hallan entre las diez con mayor índice-H del mundo, pero que no figuran entre las 100 publicaciones de mayor FI.

Adentrándonos en la respuesta a la pregunta arriba formulada, cabe reparar en algunas valoraciones que se han realizado. Por ejemplo, en un reciente y documentado artículo donde se analiza este tema,<sup>42</sup> se señala que la mezcla de revistas con repositorios "impide" la comparación bibliométrica entre aquellas y estos. Resulta difícil admitir tal afirmación: no solo no lo impide, ya que de hecho tal comparación se hace y el propio artículo da cuenta de algunos resultados que se obtienen cuando se lleva adelante, sino que no se aprecia nítidamente cuál es el problema que supone hacerlo. Que tal práctica "haga saltar por los aires el axioma" que la proscribe en el ámbito de TR, como señalan los autores, no es un argumento en sí mismo, sino un criterio autorreferencial: si un procedimiento se crea para suplir o complementar otro, no puede restársele valor por el solo hecho de que procede de manera diferente al que se considera insuficiente.

La mezcla de revistas y repositorios es algo ciertamente heterodoxo. Pero no por eso es pernicioso a los efectos de medir el impacto. Solo a modo de ilustración, cabe consignar que el hecho de que GS incluya una base de preprints de física como *ArXiv.org*, la cual aparece con un índice-H muy elevado, similar al que alcanzan *New England Journal of Medicine*, *Science* o *The Lancet*, es a mi juicio una verdadera fortaleza de este sistema. *ArXiv.org* es un proyecto "espectacularmente exitoso donde se comparten resultados bien desarrollados",<sup>43</sup> un marco de referencia cotidiano para los físicos, quienes suelen colocar sus trabajos allí a la vez que los envían a las revistas de su especialidad. Y lo cierto es que, si ostenta ese valor cimero del índice-H se debe a que tiene un impacto real y tangible, lo cual hace que la objeción de que sea incluido se torne más doctrinaria que racional.

En el trabajo que venía comentando se afirma que "es la publicación científica en una revista tras la revisión por pares lo que certifica la naturaleza científica de un trabajo".<sup>42</sup> La certifica, pero no la garantiza. Hay muchos ejemplos, pero basta uno. En la demostración de la naturaleza fraudulenta de varios de los artículos que aparecieron en revistas arbitradas de alto FI, avalando el antidiabético Avandia, en virtud de la cual *GlaxoSmithKline* acaba de pagar un gravamen multimillonario, intervino medularmente el análisis realizado con estudios solo ubicados en repositorios.<sup>44</sup> Por otra parte, está el problema de la naturaleza selectiva de los trabajos que son publicados en las revistas indexadas por TR. Recientemente, por ejemplo, se demostró que de los ensayos relacionados con antidepresivos que fueran enviados a la FDA (*Food and Drug Administration*), virtualmente todos los que arrojaron resultados "positivos" habían sido publicados en tales revistas, a la vez que casi ninguno de aquellos cuyos resultados fueron "negativos" lo fue.<sup>45</sup>

Si bien los méritos del índice-H han sido casi unánimemente reconocidos<sup>46,47</sup> el más fuerte indicio de que este indicador alcanza muy altos niveles de aceptación como medida de éxito académico es quizás el hecho de que haya sido incluido como parte del nuevo "citation report" emitido por la *Web of Science*, que gestiona la empresa *Thomson Reuters*. En este contexto, un importante elemento complementario a los efectos de valorar su eficacia viene dado por la constatación de que las revistas inscritas en la órbita del acceso abierto indexadas por *Web of Science* y/o *Scopus* tienen niveles de impacto y calidad similares a las que se mantienen en el régimen de suscripción, especialmente en el ámbito de la biomedicina.<sup>48</sup>

El índice-H no está umbilicalmente relacionado con *Google Scholar*. Puede calcularse empleando otras fuentes, tales como la propia *Web of Science*, pero lo cierto es que en la actualidad se apuesta crecientemente por calcularlo a través de dicho buscador, como opción que contempla lo que se ha publicado en diversos ámbitos y, en especial, en el del acceso abierto. El grupo de investigación australiano *Tarma Software Research* desarrolló el programa *Publish or Perish*, que se puede descargar de forma gratuita (<http://www.harzing.com/resources.htm#/pop.htm>) y que facilita la determinación del índice-H y algunas de sus variantes con el empleo de los documentos albergados en línea en este servicio de Google.

Empleando este poderoso recurso, por poner un ejemplo que resalta sus potencialidades, un análisis reciente<sup>49</sup> examinó las citas recibidas por la Revista Cubana de Información en Ciencias de la Salud (*Acimed*) a lo largo de sus 15 primeros años de existencia, en comparación con otras 9 revistas publicadas en idioma castellano sobre Bibliotecología y Ciencias de la Información. Se trata de algo difícilmente conseguible empleando otros recursos si se quiere hacer con el rigor que dicho estudio despliega.

## LA MÉTRICA DE GOOGLE ACADÉMICO

Google inauguró en 2012 un servicio oficialmente denominado como *Google Scholar Metrics* (GSM)<sup>50</sup> que permite conocer las revistas más destacadas para cada uno de diez idiomas. Operando con una métrica muy similar a la del índice-A arriba descrito y ciñéndose a la producción de los últimos cinco años, coloca en orden descendente las 100 revistas más influyentes del mundo. Excluye las publicaciones con menos de 100 artículos en ese período, así como los artículos que no hayan sido citados ninguna vez.

*Google Scholar Metrics* computa el índice-H para cada publicación, identifica el núcleo-H y calcula la mediana (*H-median*) de las citaciones que han recibido los artículos que integran el núcleo.

Por ejemplo, si la revista tuviera nueve artículos en total y los números de citas que han recibido en los últimos cinco años fueran respectivamente 37, 14, 11, 7, 5, 4, 3, 2 y 2, el índice-H sería igual a  $t=5$ , ya que 5 de los 9 artículos tienen al menos 5 citas, a la vez que ninguno de los restantes 4 supera las 5 citas. En tal caso, el núcleo-H está conformado por 5 artículos (aquellos que han recibido 37, 14, 11, 7 y 5 citas respectivamente). La mediana de esos 5 valores, 11 en este ejemplo, es un indicador que este servicio agrega complementariamente al valor del índice-H, que es el que realmente emplea para establecer el *ranking* que publica.

Nótese que esta maniobra contempla el monto de citas acopiado por los artículos que integran el núcleo-H. Por ejemplo, si otra revista tuviera también 9 artículos, pero sus números de citas fueran 10, 9, 8, 7, 5, 4, 3, 2 y 2, entonces el índice-H sería  $t=5$ , igual que la anterior, pero la mediana-h sería igual a 8, quedando así por debajo de la primera.

GSM incluye los artículos de revistas que se hallan en sitios Web indexados por *Google Scholar*, cualquiera sea el idioma en que hayan sido escritos, así como artículos presentados en conferencias y preprints de un pequeño número de fuentes identificadas manualmente por sus gestores; excluye patentes, libros y disertaciones. Debe decirse que Google no es particularmente explícito en cuanto a los criterios de elección para su indexación, lo cual no deja de ser una deficiencia que debería superarse en el futuro. Sin embargo, una revisión empírica permite apreciar que dichos criterios son ampliamente inclusivos. Particularmente, el hecho de que figuren los más importantes repositorios lo pone de manifiesto.

El listado que genera GSM contempla todo tipo de revistas científicas (publicaciones biosanitarias, pero también de economía, sociología, geografía, psicología y otras disciplinas). Resulta estimulante para Cuba constatar que cuando se aplica este recurso para las revistas que se publican en castellano, en el presente (agosto de 2012) figuran dos revistas cubanas dentro de la lista de las 100 más destacadas del mundo. Ambas resultan ser del ámbito de la salud: la *Revista Cubana de Salud Pública* (lugar 30º) y la *Revista Cubana de Información en Ciencias de la Salud*, ACIMED (lugar 47º).

## PUNTUALIZACIONES FINALES Y ÁREAS PARA EL ANÁLISIS

Parece no haber dudas de que el mundo cuantitativo está en permanente movimiento. A mí juicio está claro que los recursos convencionales vertebrados en

torno a lo que fuera el ISI y hoy se atrincheran en las reglas que impone *Thomson Reuters* son insuficientes para facilitar el intercambio entre científicos de habla no inglesa, medir su impacto con raseros transparentes, y colocar su obra en espacios con la visibilidad que pudiera corresponderle por derecho propio.

Recientemente, el proyecto *Scielo* (*SCientific Electronic Library Online*), promotor de la literatura científica iberoamericana, ha alcanzado un acuerdo con *Thomson Reuters*, de manera que buena parte de las revistas incorporadas dentro de dicho proyecto quedarían indexadas en el entorno de la transnacional. ¿Una buena noticia? En principio, hay que responder afirmativamente, ya que "nuestra" literatura, así como sus centros y sus países, alcanzarán sin duda mayor visibilidad. Pero considero que es un asunto polémico, pues al mismo tiempo *Scielo* está convalidando la lógica empresarial "y por tanto técnica y política" de la transnacional. Se trata de un acuerdo tecnológico, pero también comercial (*Scielo* y cada uno de los países integrados en ese sistema pagan en metálico por esto), cuyos detalles no han sido hechos públicos. Corresponde estudiar más detenidamente el asunto, evaluando, por ejemplo, qué implicaciones tiene y qué restricciones supone a los efectos de la creación, defensa, y consolidación de sistemas no dependientes de vértices de poder cuya historia no se ha caracterizado precisamente por anteponer la objetividad y la imparcialidad a sus intereses comerciales.

Ni Google académico, ni el índice-H, ni la conjunción de ambos recursos conforman una solución óptima para el complejo desafío de medir el impacto de una revista. De hecho, el índice comparte algunos de los defectos se le han atribuido al factor de impacto. Por ejemplo, la sensibilidad que padece ante la posibilidad de que las revistas se autociten artificialmente.<sup>51</sup> Tampoco dan solución definitiva al siempre preterido asunto de colocar a la ciencia iberoamericana en su justa posición. Pero, desde mi punto de vista y mi experiencia, abren una avenida interesante y novedosa. Y considero que aportan un mecanismo más útil en el proceso de descolonización científico-técnica que ningún otro recurso a nuestro alcance.

## REFERENCIAS BIBLIOGRÁFICAS

1. Garfield E, Sher IH. New factors in the evaluation of scientific literature through citation indexing. *American Documentation*. 1963;14(3):195-201.
2. Martyn J, Gilchrist A. An evaluation of British scientific journals (Aslib Occasional Publication, no. 1). London: Aslib, 1968.
3. Garfield E. Citation analysis as a tool in journal evaluation. *Science*. 1972; 178(4060):4719.
4. Archambault É, Larivière V. History of journal impact factor: contingencies and consequences, *Scientometrics*. 2009;79(3):639-53.
5. Seglen PO. Why the impact factor of journals should not be use for evaluating research. *British Medical Journal*. 1977;314:497-507.
6. Moed HF, Van Leeuwen TN. Improving the accuracy of Institute for Scientific Informations journal impact factors. *Journal of the American Society for Information Science*. 1995;46(6):461-7.
7. Glänzel W, Moed HF. Journal impact measures in bibliometric research. *Scientometrics*. 2002;53(2):171-93.

8. Aréchaga JM. Las revistas profesionales como claves para el desarrollo de la ciencia, la medicina y la tecnología en España. *Panace@*. 2005;4(19).
9. Rossner M, Van Epps H, Hill E. Show me the data. *J Cell Biol*. 2007;179(6):10911092.
10. Smith R. Journal accused of manipulating impact factor. *British Medical Journal*. 1997;314(7079):463.
11. Metze K. Bureaucrats, researchers, editors, and the impact factor - a vicious circle that is detrimental to science. *Clinics*. 2010;65(10):937-40.
12. Hernán MA. Epidemiologists (of all people) should question journal impact factors. *Epidemiology*. 2008;19:3668.
13. PLoS Medicine Editors. The impact factor game: It is time to find a better way to assess the scientific literature. *PLoS Medicine*; 2006;3:291.
14. Brumback RA. Impact Factor: Let's Be Unreasonable! *Epidemiology*. 2009;20(6):932-3.
15. Garfield E. The history and meaning of the Journal Impact Factor. *J Am Med Associat*. 2007;295(1):90-3.
16. Arencibia JR, Moya F. La evaluación de la investigación científica: una aproximación desde la perspectiva cuantitativa. *Acimed*. 2008;17(4) [citado 14 de septiembre de 2012]. Disponible en [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S102494352008000400004&lng=es&nrm=iso](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S102494352008000400004&lng=es&nrm=iso) >.
17. Silva LC. Un indicador obsoleto pero lamentablemente vigente. *Enfermería Clínica*. 2011;21:59.
18. Gugliotta G. The genius index: One scientist's crusade to rewrite reputation rules. *Wired Magazine*. 2009;17(6):92-5.
19. Hirsch JE. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*. 2005;102(46):1656972.
20. Vieira ES, Gomes JA. Research impact indicator for institutions. *J Informetr*. 2010;4(4):581-90.
21. Egghe L. Theory and practice of the g-index, *Scientometrics*. 2006;69(1):131-52.
22. Rousseau R. New developments related to the Hirsch index. *Science Focus*. 2006;1(4):235.
23. Jin B. H-index: an evaluation indicator proposed by scientist. *Science Focus*. 2006;1(1):8-9.
24. Kosmulski M. A new Hirsch-type index saves time and works equally well as the original h-index. *ISSI Newsletter*. 2006;2(3):4-6.
25. Sidiropoulos A, Katsaros D, Manolopoulos Y. Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics*. 2007;72(2):253-80.

26. Jin BH, Liang LM, Rousseau R, Egghe L. The R- and AR-indices: Complementing the h-index. *Chinese Science Bulletin*. 2007;52(6):855-63.
27. Anderson TR, Hankin KSH, Killworth PD Beyond the Durfee square: Enhancing the h-index to score total publication output. *Scientometrics*. 2008; 76(3):577-88.
28. Zhang CT. The e-Index, Complementing the h-Index for excess citations. *PLoS ONE*. 2009;4(5):5429.
29. García MA. A multidimensional extension to Hirsch's h-index. *Scientometrics*. 2009;81(3):779-85.
30. Brown RJC. A simple method for excluding self-citation from the h-index: the b-index. *Online Information Review*. 2009;33(6):1129-36.
31. Cabrerizo FJ, Alonso S, Herrera E, Herrera F. q2-Index: Quantitative and Qualitative Evaluation Based on the Number and Impact of Papers in the Hirsch Core. *J Informetr*. 2009;4(1):23-8.
32. Rousseau R, Ye FY. A proposal for a dynamic h-type index. *J Am Soc Inform Sci Technol*. 2008;59(11):1853-5.
33. Glanzel W, Schubert A. Hirsch-type characteristics of the tail of distributions. The generalised h-index. *J Informetr*. 2009;4(1)118-23.
34. Kosmulski M. Hirsch-type index of international recognition. *J Informetr*. 2010;4(3):351-7.
35. Bornmann L, Mutz R, Daniel HD. The h index research output measurement: Two approaches to enhance its accuracy. *J Informetr*. 2010;4(3):407-14.
36. Alonso S, Cabrerizo FJ, Herrera E, Herrera F. hg-index: A new index to characterize the scientific output of researchers based on the h- and g- indices. *Scientometrics*. 2010;82(2):391-400.
37. Namazi MR, Fallahzadeh MK. N-index: A novel and easily-calculable parameter for comparison of researchers working in different scientific fields. *Ind J Dermatol, Venereol Leprol*. 2010;76(3):229-30.
38. Harzing AW, van der Wal R. Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*. 2008;8(1):6173.
39. Bar-Ilan J. Which h-index? - A comparison of WoS, Scopus and Google Scholar. *Scientometrics*. 2008;74(2):257-71.
40. Jacso P. The plausibility of computing the h-index of scholarly productivity and impact using reference-enhanced databases. *Online Information Review*. 2008;32(2):266-83.
41. Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals. *Scientometrics*. 2006;69(1):16973.
42. Delgado E, Cabezas A. Google Scholar Metrics: una herramienta poco fiable para la evaluación de revistas científicas. *El profesional de la información*. 2012;21(4):419-27.

43. Shrager J, Billman D, Convertino G, Massar JP, Pirolli P. Soccer science and the Bayes community: Exploring the cognitive implications of modern scientific communication. *Topics in Cognitive Science*. 2009;2(1):53-72.
44. Silva LC. La industria farmacéutica y los obstáculos para el flujo oportuno de información: consecuencias para la salud pública. *Rev Cubana Sal Públ*. 2011;37(Supl. 5):631-43.
45. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*. 2008;358:25260.
46. Baldock C, Ma RMS, Orton CG. The h-index is the best measure of a scientist's research productivity. *Medical Physics*. 2009;36(4):10435.
47. Glänzel W. On the opportunities and limitations of the H-index, *Science Focus*. 2006;1(1):10-11.
48. Björk BC, Solomon D. Open access versus subscription journals: a comparison of scientific impact. *BMC Medicine*. 2012;10:73.
49. Arencibia JR. Acimed en Scholar Google: un análisis de citas de la Revista Cubana de los Profesionales de la Información y la Comunicación en la Salud [Internet]. ACIMED. 2008;18(1) [citado 5 de septiembre de 2012]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94352008000700003&lng=es&nrm=iso](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352008000700003&lng=es&nrm=iso).
50. Google Scholar Metrics [Internet]. 2012 [citado 20 de agosto de 2012]. Disponible en: [http://scholar.google.com/citations?view\\_op=top\\_venues&hl=en&vq=es](http://scholar.google.com/citations?view_op=top_venues&hl=en&vq=es)
51. Testa J. Playing the system puts self-citation's impact under review. *Nature*. 2008;455:729.

Recibido: 5 de julio de 2012.

Aprobado: 23 de julio de 2012.

Dr. *Luis Carlos Silva Ayçaguer*. Centro Nacional de Información de Ciencias Médicas-Infomed. Calle 27 entre M y N, Vedado, La Habana, Cuba. Correo electrónico: [lsilva@infomed.sld.cu](mailto:lsilva@infomed.sld.cu)

---

\* Formalmente, si los trabajos han sido objeto de  $x(1), x(2), \dots, x(N_p)$  citas, donde  $x(i) \geq x(i+1)$ , el valor del

índice es igual al mayor entero  $g$  para el cual se cumple que  $\sum_{i=1}^g x(i) \geq g^2$ .