



PROPUESTA DE NUEVO ÍNDICE BIBLIOMÉTRICO Y SU REPRESENTACIÓN POLAR: EL ÍNDICE DE ALCANCE INTEGRAL

Autores: Annier Jesús Fajardo Quesada^{1*}, René Herrero Pacheco²

¹Estudiante de Tercer año de Medicina, Facultad de Ciencias de Bayamo.

²Estudiante de Quinto año de Medicina, Facultad de Ciencias de Bayamo.

Universidad de Ciencias Médicas de Granma. Granma, Cuba

*annierfq01@gmail.com

RESUMEN

Introducción: los indicadores bibliométricos son variables matemáticas que permiten evaluar la calidad de las publicaciones y sus investigaciones derivadas. Categorizar a los autores y sus instituciones según la calidad de sus investigaciones por índices integrales se ha vuelto una necesidad, para la asignación de los recursos, contratación de personas o la concesión de premios y distinciones. **Objetivo:** crear un índice que evalúe de manera integral las publicaciones de investigaciones científicas. **Método:** se realizó una investigación de desarrollo e innovación tecnológica en el mes de febrero del 2023. La creación del índice se llevó a cabo en tres etapas, búsqueda de información sobre los índices existentes y sus bases teóricas de cálculo, creación de la fórmula y representación polar asociada y comprobación de la fórmula. **Resultados:** existió una correcta correlación entre él y los demás índices actualmente empleados. La comprobación informática demostró acertada elaboración de la fórmula y refuta el criterio positivo de los evaluadores. **Conclusiones:** el índice de alcance integral constituye un valor confiable para evaluar la producción científica, que muestra estrecha relación con otros índices actualmente aceptados y cuenta con representación gráfica que determina su cálculo.

Palabras Clave: Bibliometría; Gráfico; Índice; Literatura; Matemática; Tecnología.



INTRODUCCIÓN

La bibliometría es el análisis de la información publicada (libros, artículos de revistas, conjuntos de datos, blogs, etc.) y sus metadatos relacionados (resúmenes, palabras clave, citas, etc.) utilizando estadísticas para describir o mostrar las relaciones entre los trabajos publicados. La bibliometría se basa en la suposición de que la producción académica de un campo se captura en la literatura publicada⁽¹⁾.

La investigación bibliométrica ha experimentado un notable crecimiento durante las últimas dos décadas. Por ejemplo, Donthu et al. (2021) informaron que la investigación bibliométrica en las áreas temáticas de negocios, administración y contabilidad ; economía, econometría y finanzas ; y las ciencias sociales , según la clasificación de Scopus, promedió 1021 publicaciones por año en la última década (2011-2020) y alcanzó 1950 publicaciones en 2020 en comparación con 170 publicaciones en 2005⁽²⁾.

Para expresar cuantitativamente las características bibliográficas de un documento o conjunto de documentos se utilizan indicadores bibliométricos (IB). Estos indicadores son datos numéricos que representan distintas características de la actividad científica vinculadas tanto a la producción como al uso de la información. Los IB se calculan de forma objetiva y reproducible a partir de un gran volumen de datos disponibles en bases de datos de referencia internacional⁽³⁾. Plantear o formular adecuadamente un IB requiere de una matemática aplicada con bases teóricas muy bien justificadas.

La práctica de las matemáticas implica descubrir patrones y utilizarlos para formular y probar conjeturas, lo que da como resultado teoremas. Desde la década de 1960, los matemáticos han utilizado computadoras para ayudar en el descubrimiento de patrones y la formulación de conjeturas, por lo que trabajar la matemática paralelamente a la matemática constituye un apoyo vital en la comprobación de patrones y/o teorías⁽⁴⁾.



Las fórmulas que calculan índices bibliométricos se enfocan en caracterizar la productividad científica en determinados aspectos como son la productividad, transitoriedad, alcance, etc. que muchas veces tienen carácter objetivo, sin embargo, evaluar de forma integradora dicha productividad se vuelve una necesidad, no solo por el creciente auge de las publicaciones sino porque un solo aspecto no determina calidad y la subjetividad que puede abarcar dicho concepto hace difícil establecer escalas. Por tal motivo se redacta la presente investigación con el **objetivo** de crear un índice que evalúe de manera integral las publicaciones de investigaciones científicas.

MÉTODOS

Se realizó una investigación de desarrollo tecnológico o aplicado en diferentes universidades de Cuba durante los meses de noviembre del 2022 y enero del 2023. La creación y validación del índice y su representación se realizó en 5 etapas:

- 1- Búsqueda de información sobre los índices más empleados y sus variables determinantes
- 2- Aislamiento de las variables determinantes
- 3- Asociación de las variables en representación polar
- 4- Creación de la fórmula
- 5- Comprobación matemática y rectificación de la fórmula

Se realizó una revisión bibliográfica exhaustiva usando de manera automatizada el motor de búsqueda Google Académico donde las principales bases de datos utilizadas fueron PubMed, SCOPUS y WoS. La búsqueda no hizo limitación lingüística o geográfica, aunque se seleccionaron las publicaciones más recientes y actualizadas. Las estrategias de búsqueda incluyeron los términos: indicadores bibliométricos, bibliometría, cienciometría, fórmula matemática, índice bibliométrico y otros términos afines al tema de la revisión tanto en español como sus traducciones al inglés. Se utilizaron como operadores lógicos AND y OR en estos términos para aumentar la especificidad de los resultados.



Luego de obtenidos los índices bibliométricos más empleados se aislaron sus variables determinantes, o sea, los datos numéricos que formaban parte de la ecuación de cálculo de dicho índice. Se hizo una lista sin colisiones para su posterior análisis.

Mediante un plano polar (circular) se vincularon las variables determinantes, con basamento en la teoría empírica obtenida de la búsqueda inicial y con la condición de mantener la proporcionalidad entre las fórmulas. En este proceso se le dio valor a unas variables sobre otras según el nivel máximo y mínimo que pueden alcanzar, así como por los resultados de las investigaciones realizadas del área de las bibliometrías que se encontraron en la búsqueda inicial.

A partir de la representación polar se realizó una ecuación matemática dependiente del área que de las variables representadas en el plano. La fórmula fue puesta a prueba con un ejemplo hipotético (caso de prueba). Los sesgos encontrados determinaban el retroceso al paso anterior de creación de la fórmula, e incluso de readaptación de la representación polar como consecuencia de la modificación.

El gráfico se generó programando un Plot simple con Python 3.11 y el módulo externo Matplotlib.

RESULTADOS

Entre los índices bibliométricos más empleados en la actualidad se encuentran: índice h, índice de colaboración de autores, índice de multiautoría, índice de aislamiento, índice de Price, índice de inmediatez, factor de impacto, índice de citas, índice de autocitación, índice de visibilidad, índice de Osk. Las variables determinantes principales fueron: total de citas (Ct), cantidad de artículos(N), cantidad de artículos originales(O), y la colaboración de los autores(A). La proporcionalidad de ellos y sus variables determinantes se muestra en la tabla 1, donde se omitieron algunas variables por su bajo nivel de ocurrencia.



Tabla 1. Relación de proporcionalidad entre los índices bibliométricos y sus variables determinantes de cálculo.

Índice Bibliométrico	Variables determinantes			
	C _t	N	O	A
H	D			
Colaboración de autores		I		D
Multiautoría				*
Aislamiento	D			
Inmediatez	D	I		
Impacto	D	I		
Citas	D	D		
Autocitación	*	*		
Visibilidad	D			
Productividad de los autores		D		
Vida media	D			
Obsolecencia	D			
Inmediatez	D	I		
Osk			D	
Price		*		

D: Directamente proporcional **I:** Inversamente proporcional **•:** Depende de estados particulares de la variable

El número de citas, de artículos originales, de artículos con colaboración multistitucional tendió a ser directamente proporcional y la cantidad de artículos inversamente proporcional.

Debido a la importancia que juegan los resultados y su alcance se decide incluirlo como variable a tener en cuenta en la propuesta de índice. El alcance de los resultados (R) fue tratado como la propiedad de artículos y su método de tener un alcance universal, independientemente de los criterios de selección exclusión, o sea, aquellos estudios cuyos resultados no se basen en grupos dependientes de



su localización geográfica, sino que abarquen una población distribuida en todo el mundo. De igual manera se consideró el promedio de citas por artículo, para darle valor a la calidad sobre la cantidad (densidad de citas como criterio de calidad grupal).

Para poder representar en un plano polar adecuadamente los valores sin que los valores extremos deformen el cuadrilátero y alteren el cálculo de área se llevó a por ciento las variables O, I y A con base N. C (promedio de citas por artículo) por su amplia variabilidad se llevó a una escala logarítmica (en base 2) por niveles (Tabla 2) y su calificación se basó en esto (el 50 % del valor del grupo para reducir el tamaño de los datos y mejorar su proporcionalidad) y en el porcentaje que representaba su cantidad dentro del propio grupo.

Tabla 2. Puntuación por grupos según promedio de citas por artículo

Rango de los grupos	Grupos (G)	Puntuación (GP)
{0} U (0 – 2)	0	0.0
[2-4)	1	0.5
[4 – 8)	2	1.0
[8 – 16)	3	1.5
[2 ^G – 2 ^{G+1})	G	G/2

El grupo para una cantidad C de como promedio de citas por publicación está entonces determinado por la ecuación:

$$G = \lfloor \log_2 C \rfloor$$

La cantidad de citas que abarca el grupo (rango del grupo o G_r) se determina por:

$$G_r = 2^{G+1} - 2^G = 2^G$$

Luego la fracción de citas dentro de su grupo se calcula:

$$FC = (C - 2^G) / 2^G = C / 2^G - 1$$

De este modo los determinantes de cálculo serían:



Fracción de alcance de los resultados (FR): R/N

Fracción de citación (FC): $C / 2^{\lfloor \log_2 C \rfloor} - 1$

Fracción de originalidad (FO): O/N

Fracción de colaboración institucional (CI): C/N

Trabajando con un ejemplo hipotético de $C_t = 1000$, $N = 100$, $FR = 60$, $FO = 70$ y $CI = 90$ se obtiene la siguiente representación polar representada en la figura 1.

Figura 1. Representación polar para $C_t = 1000$, $N = 100$, $O = 60$, $I = 70$ y $A = 90$.



El eje de las abscisas estaría determinado por variables relacionadas a la conexión entre autores e investigaciones; y el eje de las ordenadas por variables relacionadas al aporte científico, dependientes de las investigaciones propiamente.



Entonces el valor final del índice del grupo de artículos se calcula teniendo en cuenta la puntuación del grupo al que pertenece por sus citaciones y el área de su polígono de representación (AP).

$$AP = (FO + FR) * FC / 2 + (FO + FR) * CI / 2 = (FO + FR) * (FC + CI)$$

Luego el índice de alcance integral se calcula:

$$IAI = PG + AP$$

$$\begin{aligned} IAI &= G / 2 + (FO + FR) * (FC + CI) \\ &= \lfloor \log_2 C \rfloor / 2 + (O / N + R / N) * (C / 2^{\lfloor \log_2 C \rfloor} - 1 + C / N) \\ &= \lfloor \log_2 C \rfloor / 2 + (O + R) / N * (C / 2^{\lfloor \log_2 C \rfloor} - 1 + C / N) \end{aligned}$$

Para el ejemplo hipotético analizado $IAI = 5.86$

Una alternativa que regule la sensibilidad de la escala logarítmica con base b sería:

$$IAI = \lfloor \log_b C \rfloor / 2 + (O + R) / N * (C / b^{\lfloor \log_b C \rfloor} - 1 + C / N)$$

DISCUSIÓN

La subjetividad es omnipresente e implica perspectivas que van desde lo intrapersonal hasta lo intercultural⁽⁵⁾. Cuando se trata de medir criterios de cantidad esta subjetividad se hace muy poco marcada pero cuando lo evaluado es la calidad depende mucho de su evaluador y sus parámetros individuales. La presente propuesta de índice se basa en criterios de calidad sobre cantidad ⁽⁶⁾ por lo que en todo momento se debe tener en cuenta que no se pretende alcanzar una exactitud sino acercarse a un indicador lo más integral posible, que por usarse como base la teoría proveniente de otras fórmulas aceptadas actualmente se mantiene una proporcionalidad directa entre este y dichas fórmulas.

Flores Fernández et al⁽⁷⁾ dividen a los indicadores bibliométricos en personales, de productividad, de citación, de contenido y metodológicos. Bajo esta clasificación se engloban la mayoría de los IB, pero otros, pertenecen simultáneamente a varias clasificaciones como es el caso del índice de Osk y el



índice propuesto en la presente investigación. En este sentido otra división podría ser: índices específicos (los de la clasificación de Flores Fernández) e índices integrales (donde se encuentra el índice de Osk).

El índice de Osk pretende dar solución a algunos de los problemas que presentan otros índices, tales como no medir todos los factores que intervienen en el proceso editorial de una revista científica. Los componentes que participan en la ecuación del Índice de Osk presentan unos valores acordes con la jerarquía que ocupan en las revistas, entre estos se encuentran artículos de investigación científica, revisión, ediciones, citas, etc., cada uno de los valores que intervienen en la ecuación son aditivos y/o multiplicativos, ya que todo lo que se presenta en una revista es producto del esfuerzo de un equipo editorial; por último, este índice se caracteriza por su flexibilidad, permitiendo la adición o eliminación de valores de acuerdo a las características del objeto de estudio que se desea evaluar⁽⁸⁾.

Haciendo una comparación del índice de Osk con el IAI, por sus objetivos similares, se evidencia como ambos evalúan la tipología como factor determinante, y dentro de ella se les da más peso a los artículos originales. El IAI a diferencia del IO no toma en cuenta el resto de tipologías por considerar únicamente a los artículos de investigación original, que son los que aportan conocimiento nuevo a la comunidad científica⁽⁹⁾, esto les da una menor sensibilidad a los cambios de tipología, pero da un resultado más enfocado en los aportes. Algo similar ocurre con el nivel de colaboración donde en IO se sensibiliza a identificar la colaboración local, nacional e internacional.

El IAI le da igual valor a la colaboración, alcance de los resultados y a la cantidad de originales, IO le da más valor a la colaboración. Esto demuestra la subjetividad que pueden llegar a tener las escalas de calidad. El IAI también le da un fuerte valor a la cantidad de citas, que es tan importante que gran parte de los indicadores más usados la contienen como factor de cálculo⁽⁷⁾.



Como la cantidad de citas puede ser muy irregular, y va en dependencia de la serie de artículos que se evalúen, tener una “llave de regulación de sensibilidad a las citas” es de mucha utilidad como en las bibliometrías de trabajos publicados en sitios locales donde la tasa de citación es realmente baja ⁽¹⁰⁾. Esta se representó en la fórmula como b.

Por cuestiones de aprendizaje y visibilidad muchos de los procesos son llevados a sus representaciones gráficas, siempre que se pueda, porque a veces constituye todo un reto vincular los mismos a imágenes coherentes y sencillas⁽¹¹⁾. La representación polar del IAI no es solo un implemento decorativo, pues les da cierta dependencia a las variables y la demuestra gráficamente. Es por esto que una mayor o variación del IAI dependerá simultáneamente de los valores representados en las abscisas y en las ordenadas. En la actualidad hay índices sofisticados que cuentan con su propia representación gráfica, como son los indicadores basados en mapas y grafos como el caso de los grafos de coocurrencia de términos de WoS Viewer ⁽¹²⁾.

Con el auge de las tecnologías la propiedad de los procesos de ser automatizados les da la posibilidad de una mayor generalización, alcance, productividad, costos, velocidad y reducción de errores ⁽¹³⁾. Teniendo en cuenta que las fórmulas son procesos matemáticos para obtener un valor determinado, prácticamente todas las fórmulas pueden ser automatizadas, y es el caso del IAI, lo que constituye uno de sus puntos fuertes.

CONCLUSIONES

El índice de alcance integral, a pesar de su marcada subjetividad característica de los índices que evalúan calidad, constituye un valor confiable para evaluar la producción científica de autores, instituciones y revistas; que muestra estrecha relación con algunos índices actualmente aceptados y cuenta con representación gráfica que deter



REFERENCIAS BIBLIOGRÁFICAS

1. Ninkov A, Frank JR, Maggio LA. Bibliometrics: Methods for studying academic publishing. *Perspect Med Educ* [Internet]. 2022 [citado 29 de mayo de 2023];11(3):173-6. Disponible en: <https://link.springer.com/article/10.1007/s40037-021-00695-4>
2. Mukherjee D, Lim WM, Kumar S, Donthu N. Guidelines for advancing theory and practice through bibliometric research. *J Bus Res* [Internet]. 2022 [citado 29 de mayo de 2023];148:101-15. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S0148296322003824>
3. García-Villar C, García-Santos JM. Bibliometric indicators to evaluate scientific activity. *Radiol (English Ed)* [Internet]. 2021 [citado 29 de mayo de 2023];63(3):228-35. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S2173510721000549>
4. Davies A, Veličković P, Buesing L, Blackwell S, Zheng D, Tomašev N, et al. Advancing mathematics by guiding human intuition with AI. *Nature* [Internet]. 2021 [citado 29 de mayo de 2023];600(7887):70-4. Disponible en: <https://www.nature.com/articles/s41586-021-04086-x>
5. Brown SR. Subjectivity in the Human Sciences. *Psychol Rec* [Internet]. 2019 [citado 29 de mayo de 2023];69(4):565-79. Disponible en: <https://link.springer.com/article/10.1007/s40732-019-00354-5>
6. Quesada AJF. Contaminación de la línea investigativa. El precio de la cantidad sobre la calidad. *Rev Cubana Med* [Internet]. 2023 [citado 29 de mayo de 2023];62(1). Disponible en: <https://revmedicina.sld.cu/index.php/med/article/view/3150>
7. Flores-Fernández C, Aguilera-Eguía R, Flores-Fernández C, Aguilera-Eguía R. Indicadores bibliométricos y su importancia en la investigación clínica. ¿Por qué conocerlos? *Rev la Soc Española del Dolor* [Internet]. 2019 [citado 29 de mayo de 2023];26(5):315-6. Disponible en: https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1134-80462019000500012&lng=es&nrm=iso&tlng=es
8. Pérez-Anaya O. Índice de Osk: una nueva medición bibliométrica para las revistas científicas. *Rev Española Doc Científica* [Internet]. 2017 [citado 29 de mayo de 2023];41(1):1-10. Disponible en: <https://www.rdc.es/revista/ver/10.3916/revista.41.1.001>



- mayo de 2023];40(2):e174-e174. Disponible en:
<https://redc.revistas.csic.es/index.php/redc/article/view/978/1499>
9. Fajardo Quesada AJ, Calas Torres JJ, Rego Rodríguez FA, Otaño Castillo R, Montiel Alfonso MA. Producción científica cubana sobre Dengue en Scopus en el periodo 2011-2021. Univ Médica Pinareña [Internet]. 2022 [citado 30 de mayo de 2023];18(2):e892. Disponible en:
<https://revgaleno.sld.cu/index.php/ump/article/view/892>
10. Akgülle AH, Şirin E. Publication rates and characteristics of abstracts presented in the first National Pediatric Orthopedics Congress: a bibliometric study. Kocaeli Med J [Internet]. 2020 [citado 30 de mayo de 2023];9(3):75-8. Disponible en: <https://kocaelimj.org/eng/jvi.aspx?un=KTD-15931&volume=9&issue=3>
11. Chen F, Wang YC, Wang B, Kuo CCJ. Graph representation learning: a survey. APSIPA Trans Signal Inf Process [Internet]. 2020 [citado 30 de mayo de 2023];9:e15. Disponible en: <https://www.cambridge.org/core/journals/apsipa-transactions-on-signal-and-information-processing/article/graph-representation-learning-a-survey/23B9870F91F7E6DA14784959A9BC9E7A>
12. Sindhu P, Bharti K. Mapping customer experience: a taxonomical study using bibliometric visualization. VINE J Inf Knowl Manag Syst [Internet]. 2021 [citado 30 de mayo de 2023];51(4):592-617. Disponible en: <https://www.emerald.com/insight/content/doi/10.1108/VJIKMS-11-2019-0178/full/html>
13. Tejada DMR, Navarro IJN, Ibarra CHO. Lineamientos para la Automatización de Robótica de Procesos. Rev CIES Escolme [Internet]. 2020 [citado 30 de mayo de 2023];11(1):143-58. Disponible en: <http://revista.escolme.edu.co/index.php/cies/article/view/286>

Los autores certifican la autenticidad de la autoría declarada, así como la originalidad del texto.