

## Investigación científica en el campo de las ataxias espinocerebelosas: una mirada bibliométrica

Scientific Research in The Field of Spinocerebellar Ataxias: A Bibliometric  
View

Miguel Cruz Ramírez<sup>1</sup> <https://orcid.org/0000-0002-1697-1624>

Roberto Rodríguez-Labrada<sup>2</sup> <https://orcid.org/0000-0003-3193-7683>

Luis Clodovaldo Velázquez-Pérez<sup>3\*</sup> <https://orcid.org/0000-0003-1628-2703>

<sup>1</sup>Universidad de Holguín. Holguín, Cuba.

<sup>2</sup>Centro de Neurociencias de Cuba. La Habana, Cuba.

<sup>3</sup>Academia de Ciencias de Cuba. La Habana, Cuba.

\*Autor para la correspondencia: [luisvelazquez@ceniai.inf.cu](mailto:luisvelazquez@ceniai.inf.cu)

### RESUMEN

El objetivo del presente estudio consistió en caracterizar la producción científica relacionada con el campo de las ataxias espinocerebelosas, las cuales constituyen enfermedades neurodegenerativas, manifestadas por cuadros clínicos progresivos e invalidantes. La investigación es de tipo censal-documental y recupera metadatos de Scopus, correspondientes a 5654 investigaciones relacionadas con este problema de salud, durante el período 1961-2020. El procesamiento explora las principales características bibliométricas de los documentos publicados, el ritmo de crecimiento, la paternidad de las obras, el impacto por índice de citas, así como las redes de

colaboración y la estructura que sigue el flujo del conocimiento. Se observa un despegue notable de la producción científica desde inicios de los años 90 del siglo pasado, coincidiendo con el desarrollo de investigaciones afines en el campo de la genética. También se constata un predominio en el estudio de los tipos SCA1, SCA2, SCA3, SCA6, y SCA17, donde los cuatro primeros corresponden a las ataxias de mayor prevalencia a escala mundial. El corpus documental refleja la consolidación de grupos de investigación relativamente estables, encabezados por líderes científicos y caracterizados por la ampliación sostenida de la colaboración internacional y por el trabajo colectivo e interdisciplinario. También se aprecia la tendencia hacia el aumento del número de referencias dentro de cada documento. Los mayores volúmenes productivos se concentran en países desarrollados, junto a países en vías de desarrollo donde existen elevados niveles de prevalencia en esta enfermedad.

**Palabras clave:** análisis bibliométrico; ataxia; SCA; Scopus; producción científica; enfermedades neurodegenerativas.

## ABSTRACT

The objective of the present study was to characterize the scientific production related to the field of spinocerebellar ataxias, which constitute neurodegenerative diseases, manifested by progressive and disabling clinical conditions. The research is census-documentary type and recovers metadata from Scopus, corresponding to 5,654 investigations related to this health problem, during the period 1961-2020. The processing explores the main bibliometric characteristics of the published documents, the pace of growth, the authorship of the works, the impact by citation index, as well as the collaboration networks and the structure that follows the flow of knowledge. A notable takeoff in scientific production has been observed since the beginning of the 90s of the last century, coinciding with the development of related research in the field of genetics. There is also a predominance in the study of the types SCA1, SCA2, SCA3, SCA6, and SCA17, where the first four correspond to the most prevalent ataxias worldwide. The documentary corpus reflects the consolidation of relatively stable research groups, headed by scientific leaders and characterized by the sustained expansion of international collaboration and collective and interdisciplinary work. There

is also a trend towards increasing the number of references within each document. The largest productive volumes are concentrated in developed countries, along with developing countries where there are high levels of prevalence of this disease.

**Keywords:** bibliometric analysis; ataxia; SCA; Scopus; scientific production; neurodegenerative diseases.

Recibido: 05/08/2022

Aceptado: 11/12/2023

## Introducción

Las ataxias espinocerebelosas (SCAs, del inglés *spinocerebellar ataxias*) son enfermedades autosómicas producidas por mutaciones relacionadas con la codificación anormal de proteínas.<sup>(1,2,3,4)</sup> Estos trastornos generalmente se heredan como producto de un gen dominante, proveniente de alguno de ambos progenitores y ocasionan una enfermedad neurodegenerativa, caracterizada por efectos clínicos progresivos e invalidantes.

Actualmente se conocen 48 formas moleculares diferentes de SCA y su clasificación responde al gen específico que está alterado.<sup>(5)</sup> Las SCAs más comunes son la SCA1 (enfermedad de Schut), la SCA2 (ataxia cubana), la SCA3 (enfermedad de Machado-Joseph) y la SCA6.<sup>(1)</sup> En general, los tipos SCA 9-48 son raros y consecuentemente menos estudiados. En el caso de la SCA2, Cuba posee tasas de prevalencia superiores a 20 por cada 100 mil habitantes, mayormente concentrados en la provincia de Holguín y el municipio de Báguano alcanza una prevalencia de 141,66 por cada 100 mil habitantes.<sup>(3)</sup>

Un estudio epidemiológico transversal reportado recientemente en España estimó la prevalencia de ataxias hereditarias en un 5,48 por cada 100 mil habitantes y el tipo

dominante fue la SCA3.<sup>(6)</sup> Estos ejemplos son parte de un importante problema de salud a escala mundial, razón por la cual el 25 de septiembre de cada año aparece señalado como *Día Internacional de la Ataxia* con la finalidad de crear conciencia acerca de esta afectación.

La acumulación de proteínas mal plegadas, derivadas de varias de estas mutaciones, provoca un daño gradual y sostenido en las células nerviosas que componen el cerebelo y la médula espinal. Producto de ello, el paciente manifiesta signos y síntomas, tales como la pérdida del control muscular que le impide coordinar los movimientos voluntarios, dificultades para fijar la mirada, enlentecimiento de los movimientos oculares sacádicos, trastornos de la ejecución motora del habla, como la disartria, entre otros padecimientos.<sup>(3)</sup> Las primeras señales están relacionadas con alteraciones en la velocidad, el ritmo y la precisión en el andar, que luego se extienden hacia las extremidades superiores y el tronco. Los músculos se debilitan y se atrofian con el paso del tiempo, incluso aparecen deformidades y la enfermedad culmina con el fallecimiento del paciente.

Ya que no existe cura para dichas enfermedades, los tratamientos se centran en el mejoramiento de la calidad de vida de quienes las padecen. Por tanto, la atención clínica y fisioterapéutica se enfoca desde el punto de vista multidisciplinario. Incluso el trabajo se extiende al contexto de la educación y la atención especializada se centra en las necesidades educativas especiales correspondientes.<sup>(7)</sup> Por otro lado, los esfuerzos también enfrentan retos, desde puntos de vista relacionados con el desarrollo tecnológico y la investigación científica, aspecto que resulta tangible bajo una mirada bibliométrica.<sup>(8)</sup>

Investigaciones contemporáneas sobre las ataxias han mostrado la importancia que reviste el saber científico con carácter documental. Por ejemplo, en un estudio reciente se ha descrito la variabilidad de las tasas de prevalencia en diferentes regiones de Brasil y se han identificado las principales mutaciones encontradas con tecnologías de secuenciación de última generación.<sup>(9)</sup> Asimismo, con ayuda de un estudio metaanalítico de fuentes recuperadas de Medline, la SCI-WoS y Scopus se confrontaron 22 estudios realizados en 16 países, lo cual permitió comparar la prevalencia de esta enfermedad en diferentes regiones del mundo.<sup>(10)</sup>

Por ejemplo, el tipo SCA2 mostró mayores niveles en Cantabria (España) y fue segundo en Padua (Italia), en el sureste de Noruega y en Singapur. Por su parte, el tipo SCA6 mostró la segunda mayor prevalencia en los Países Bajos y en Japón. Análogamente, con base en la información recuperada de MedLine se efectuó un estudio comparado de los tamaños de muestra, donde se concluye que la progresión de la enfermedad ocurre con mayor rapidez en pacientes con SCA1, es intermedio en pacientes con SCA2 y SCA3 y se desarrolla más lento en pacientes con SCA6.<sup>(1)</sup>

Junto a los estudios metaanalíticos, basados en fuentes científicas documentales, las investigaciones bibliométricas y cienciométricas también proveen información valiosa para campos disímiles, entre los que se incluyen la genética y la neurología. Por ejemplo, a partir de una investigación relacionada con dos enfermedades raras que generan trastornos mentales y del comportamiento (CADASIL y Síndrome de Rett) durante una década de producción científica registrada en SCI-WoS y en Pubmed se constata que, si bien existen diferencias en los patrones de investigación en ambas enfermedades, estas confluyen en el empleo sistemático de los avances de la genética.<sup>(11)</sup>

De modo similar, otra investigación explora la productividad científica de la India durante el período 2008-2017, relacionada con trastornos genéticos y recuperada a partir de una búsqueda en Scopus.<sup>(12)</sup> En este caso se observa un desarrollo sostenido y acentuado en el año 2012 con 36,27 % de tasa de crecimiento. Además, dicho estudio reveló que este campo de investigación se relaciona estrechamente con la Medicina, la Bioquímica, la Genética y la Biología Molecular.

En el marco específico de las neurociencias una investigación reporta que los mayores índices de crecimiento de la producción científica global se experimentan desde el año 2015 con tasas que oscilan entre el 0,74 y 2,25 %, sin cambios importantes en el período 1986-2015.<sup>(13)</sup> Por su parte, otros estudios constatan que la enfermedad neurodegenerativa más estudiada consiste en el mal de Alzheimer, mientras que el mal de Parkinson resulta la afección más investigada en el Sudeste Asiático.<sup>(14,15)</sup> Los estudios relacionados con la rehabilitación alcanzan pesos considerablemente bajos (peso = número de documentos × índice de impacto), que oscilan entre el 0,1 y el 7 %.<sup>(16)</sup> Particularmente, en el campo específico de las SCAs, los estudios bibliométricos son

más escasos. Por este motivo, tomando en consideración las oportunidades que la bibliometría brinda para ahondar en el conocimiento acerca del desarrollo de la ciencia, surge la motivación de la presente investigación.

El objetivo fundamental del presente estudio consistió en caracterizar la producción científica relacionada con el campo de las SCAs, con base en una muestra bibliográfica suficientemente representativa y de alta calidad, medible por índices de impacto.

## Métodos

La investigación es de naturaleza bibliométrica y se apoya en indicadores relacionados con publicaciones afines de carácter científico y académico. Las investigaciones bibliométricas, si bien no captan la actividad científica en toda su magnitud, aportan información útil para caracterizarla. Ello se ha puesto de manifiesto en estudios relacionados con las ciencias biomédicas y de la salud, las cuales han permitido analizar y comparar niveles de desarrollo científico, identificar redes de colaboración, evaluar el impacto de los resultados, analizar la cobertura y difusión del conocimiento, entre otros aspectos.<sup>(12,17,18,19,20)</sup>

Se utilizaron registros almacenados en Scopus, no solo por constituir una base de datos voluminosa de la corriente principal en la actualidad, sino también por la alta calidad de la información que indexa.<sup>(21,22,23)</sup> Para cumplir los objetivos trazados la metodología sigue dos etapas sustantivas. En primer lugar, se establece una matriz de datos suficientemente amplia y confiable, capaz de aprehender la mayor cantidad de información relacionada con investigaciones sobre SCAs (recopilación de tipo censal-documental). Vale señalar que, si bien el término SCA es de uso extendido, dada la elevada diversidad etiológica de las ataxias, en ocasiones los autores utilizan otras denominaciones como “cerebellar ataxia” y “hereditary ataxia”. En segundo lugar, se realiza un análisis de variables bibliométricas fundamentales que viabilizan la caracterización definida en el objetivo del presente estudio.

## Búsqueda y curación de los datos

El día 5 de diciembre de 2021 se realiza una búsqueda de aquellos documentos que contienen el texto “spinocerebellar ataxia”, al menos en sus respectivos títulos, resúmenes o palabras clave. Debido a restricciones técnicas del motor de búsqueda, solo es posible la descarga directa de metadatos correspondientes a no más de 2000 documentos. Por tanto, se subdividen los intervalos de búsqueda en períodos que no superen la mencionada limitación. Las sentencias lógicas utilizadas tienen la siguiente estructura, donde los  $i$ -ésimos intervalos son tres:  $[X_i, Y_i] = [1961; 2008], [2007; 2015]$  y  $[2014; 2020]$ , respectivamente: TITLE-ABS-KEY (“spinocerebellar ataxia”) AND PUBYEAR >  $X_i$  AND PUBYEAR <  $Y_i$ .

Toda la información se integra en sendos archivos en formatos CSV y BibTeX, los cuales se inspeccionan paralelamente de modo general, en busca de uniformidad nominal y estructural. Por un lado, se homogenizan los apellidos de autores, los nombres de instituciones y países, así como las siglas empleadas en algunas denominaciones, principalmente en palabras clave y en resúmenes. Por otro lado, se sustituyen caracteres latinos extendidos, como  $\acute{s}$ ,  $\grave{t}$ ,  $\bar{u}$ ,  $\grave{u}$ ,  $\acute{u}$ ,  $\acute{z}$ , para evitar conflictos en el reconocimiento automatizado. También se rellenan manualmente celdas carentes de información relevante como el año de publicación, pero recuperable de modo alternativo en Google Scholar y se establece uniformidad en el formato de campos contentivos de palabras claves y referencias. Por ejemplo, tomando ventaja de que el carácter “;” sirve de separador para delimitar los elementos en el metadato “References”, la siguiente función de Excel posibilita el cálculo puntual del número de referencias almacenadas en la celda A1, lo cual resulta práctico en muchas correcciones y chequeos primarios:

$$=IF(LEN(TRIM(A1))=0,0,LEN(TRIM(A1))-LEN(SUBSTITUTE(A1,“;”,“”)))$$

---

## Procesamiento y análisis de cada matriz de datos

La matriz en formato CSV queda estructurada por 25 metadatos estándares de Scopus que facilitan el estudio de cuatro variables fundamentales. Estas consisten en características bibliométricas de los documentos publicados, ritmo de crecimiento, paternidad de la obra e impacto por índice de citas, así como redes de colaboración y estructura que sigue el flujo del conocimiento, con base en el análisis de palabras clave y resúmenes. El chequeo primario de la información se realiza en Excel y para el análisis estadístico se implementa el lenguaje R, por intermedio de funciones habilitadas en las bibliotecas “car” y “MASS”.<sup>(24)</sup> Los recursos mencionados permiten explorar y especificar la evolución diacrónica de las publicaciones, y también delimitar concentraciones de la productividad y del índice de citas.

Por otro lado, la matriz en formato BibTeX se procesa con ayuda de la interface web “biblioshiny”, correspondiente a la biblioteca especializada “bibliometrix” del lenguaje R.<sup>(25)</sup> Los códigos abiertos de algunas funciones gráficas se adaptan a las especificidades de la presente investigación. Para lograr mayor compatibilidad, se codifica la información a formato UTF-8. Los resultados se comparan y se complementan. La minería de datos y la visualización de redes bibliométricas se despliega también con ayuda de la herramienta de software VOSviewer.<sup>(26)</sup> Todo esto facilita la identificación de estructuras latentes que permiten tipificar posibles redes de colaboración y delimitar la estructura en que fluye los contenidos esenciales de la información.

## Resultados y discusión

### Características bibliométricas de los documentos

La búsqueda por intervalos produce 1762, 1929 y 1963 resultados, respectivamente, para un total de 5654 documentos publicados durante el período 1961-2020. Si se incorporan otras denominaciones relacionadas como “cerebellar ataxia” y “hereditary ataxia”, mediante el operador lógico OR, el número de documentos asciende a 15 770,

lo cual se consideró cuantiosamente complejo de analizar, desde el punto de vista etiológico.

En la tabla 1 se sintetiza un grupo de indicadores bibliométricos generales. Los mayores problemas de normalización se concentran en nombres propios y palabras clave, lo cual ha sido reportado en otras investigaciones bibliométricas relacionadas con las ciencias biomédicas y de la salud, tanto en la WoS como en Scopus.<sup>(27,28,29)</sup> La corrección manual permitió normalizar los apellidos de 21 autores y los nombres de 17 instituciones, localizados en los deciles de mayor frecuencia. Por ejemplo, el *Hôpitaux Universitaires Pitié Salpêtrière* es también conocido como *La Salpêtrière*, aunque también los autores utilizan otras formas como: “Hop. de la Salperriere”, “Hop. de la Salpêtrière”, “Hôpital de la Salpetriere”, “Hôpital de la Satpêtrière”, “Hosp. Pitie-Salpetriere”, “Hôpital de la Salpêtrière” y “Hopital de la Salpetriere”. Este ejemplo es ilustrativo de las dificultades que acarrea una falta de normalización.

**Tabla 1** - Indicadores bibliométricos generales

	<b>Autores*</b>	<b>Países</b>	<b>Afiliaciones†</b>	<b>Referencias‡</b>	<b>Palabras clave de autor§</b>	<b>Índice de palabras clave  </b>	<b>Citas</b>
Total	17754	96	16841	216874	18704	212822	155990
Promedio por documento	6,64	1,39	3,83	41,22	5,30	40,06	31,53
Desviación típica	5,24	1,03	3,51	35,72	2,07	17,94	66,92

Leyenda: \*Se excluyen 22 documentos con ID ausentes (0,39 %); †Se excluyen 130 documentos con afiliaciones ausentes (2,30 %);

‡Se excluyen 393 documentos con referencias ausentes (6,95 %); §Se excluyen 2124 documentos carentes de palabras clave (37,57 %); ||Se excluyen 38 entradas con error y 310 valores ausentes (6,15 %).

*Fuente:* Elaboración propia.

Un total de 2124 registros (37,57 %) carece de palabras clave, las cuales están casi ausentes en el transcurso de las primeras dos décadas. A partir de 1982 la inclusión de este metadato muestra un crecimiento paulatino y está presente en tres de cada cuatro documentos durante el último decenio. La ausencia de palabras clave no solo responde

a una pérdida de información, sino al hecho de que algunas revistas no demandan este metadato, como también reportan otros autores.<sup>(28)</sup> Por ejemplo, los documentos procedentes de las revistas *Annals of Neurology*, *Archives of Neurology*, *Human Molecular Genetics*, *Journal of Biological Chemistry*, *Neurology*, *PLoS ONE* carecen de palabras clave y representan el 27,59 % de toda la información recuperada. Para compensar el análisis, también se tomó en consideración el campo "Index Keywords" durante el procesamiento estadístico.

El índice de palabras clave en Scopus proviene mayormente de tesauros, como el *Medical Subject Headings* (MeSH) para las ciencias biomédicas y de la salud, donde se proveen términos normalizados. Como ya ha sido señalado, la ventaja de utilizar palabras clave de autor reside en que los investigadores tienen un mejor conocimiento del dominio que los indexadores.<sup>(30)</sup> Sin embargo, un indexador puede tener un mejor conocimiento en la selección de términos que podrían mejorar la recuperabilidad del documento. Por tanto, el empleo combinado de ambos metadatos puede suministrar una información más confiable. En total se unificaron bajo denominaciones únicas 13 conjuntos de palabras clave equivalentes, provenientes del metadato "Author Keywords" y de elevada frecuencia absoluta, así como nueve variaciones de términos similares, resultantes de la minería de texto en el metadato "Abstract". Por ejemplo, los siguientes términos se unificaron por referir el mismo elemento conceptual: "polyglutamine disorders" = "poly-glutamine (PolyQ) disorders" = "polyglutamine diseases" = "polyglutamine (PolyQ)".

La literatura científica compilada aparece redactada en 22 idiomas, con predominio del inglés (87,83 %) y seguido por el japonés (2,44 %), el chino (1,95 %) y el español (0,99 %). Un pequeño segmento de 39 artículos aparece redactado en forma bilingüe. Aparece un total de 4104 artículos en revistas (72,59 %), 592 revisiones (10,47 %), 260 cartas al editor (4,60 %), 185 capítulos de libro (3,27 %), 112 ponencias en conferencias (1,98 %), 67 notas (1,19 %), 51 reportes breves (0,90 %), 32 notas editoriales (0,57 %), 32 informes de erratas (0,57 %), tres artículos en prensa (0,05 %), dos revisiones de conferencia (0,04 %) y un libro (0,02 %). En 213 registros (3,91 %) no se consigna el tipo de documento.

Entre revistas, actas de eventos científicos y otros medios especializados aparecen 1182 fuentes, de las cuales una decena (0,84 %) contiene cerca de la quinta parte del total de documentos (23,31 %), mientras que 1001 (84,54 %) no sobrepasan los cinco documentos cada una. La tabla 2 ilustra algunos indicadores relacionados con las diez revistas más productivas en el campo de las SCAs dentro de Scopus. Los indicadores locales se basan en la matriz de datos, lo cual ayuda a establecer comparaciones relativas a este campo específico de investigación, mientras que los globales son tomados del *ranking* SJR2020 de SCImago y basado en toda la información almacenada en Scopus.

**Tabla 2 - Ranking** de revistas que han publicado artículos en el campo de las SCAs

Nº	Revista	ISSN	País / Editorial	Indicadores locales*				Indicadores globales†		
				Total	Índice h	Índice m	Índice g	Ranking	Cuartil	Índice h
1	<i>The Cerebellum</i>	1473-4222, 1473-4230	Estados Unidos / Springer New York	219	38	1,90	56	1,42	Q1	72
2	<i>Journal of the Neurological Sciences</i>	0022-510X, 1878-5883	Países Bajos / Elsevier	161	28	0,68	38	0,96	Q2	137
3	<i>Movement Disorders</i>	0885-3185, 1531-8257	Estados Unidos / John Wiley and Sons Inc.	160	41	1,58	55	3,35	Q1	198
4	<i>Human Molecular Genetics</i>	0964-6906, 1460-2083	Reino Unido / Oxford University Press	152	66	2,28	105	2,81	Q1	276
5	<i>Neurology</i>	0028-3878, 1526-632X	Estados Unidos / Lippincott	144	50	1,00	80	2,91	Q1	364

			t Williams and Wilkins Ltd.							
6	<i>Journal of Neurology</i>	0340-5354, 1432-1459	Alemania / D. Steinkopf f-Verlag	124	32	1,07	48	1,54	Q1	136
7	<i>Brain</i>	0006-8950, 1460-2156	Reino Unido / Oxford Universit y Press	107	51	1,70	80	5,14	Q1	336
8	<i>PLoS ONE</i>	1932-6203	Estados Unidos / Public Library of Science	89	26	1,86	36	0,99	Q1	332
9	<i>Parkinsoni sm &amp; Related Disorders</i>	1353-8020, 1873-5126	Países Bajos / Elsevier BV	82	20	0,83	28	1,53	Q1	98
1 0	<i>Archives of Neurology (JAMA Neurology)</i>	2168-6149, 2168-6157	Estados Unidos / American Medical Associati on	80	35	0,92	57	5,30	Q1	231

Leyenda: \*Basados en el análisis intrínseco de la matriz de datos BibTeX, en el campo de las SCAs; †Basados en SJR (*SCImago Journal & Country Rank*), respecto a toda la base de datos Scopus.

Fuente: Elaboración propia.

En la tabla 2 predominan revistas del primer cuartil de Scopus en campos de la Biomedicina, la Neurología, la Genética o bien la Biología Molecular. *The Cerebellum* encabeza la lista con 219 documentos relacionados con estudios sobre SCAs. Sin embargo, el mayor impacto intrínseco se concentra en las publicaciones de la revista *Human Molecular Genetics* con un índice de Hirsch ( $h = 66$ ).<sup>(31)</sup> Por tanto, existen a lo

sumo 66 documentos con al menos 66 citas, correspondientes a publicaciones en este campo específico de investigación y bajo los elevados estándares de calidad que demanda Scopus. Si se toma en consideración el impacto de las citas durante un tiempo  $n$  medido en años, también puede constatarse que esta última revista alcanza el mayor valor ( $m = 2,28$ ,  $h \sim mn$ ). Ello ocurre no solo dentro de este grupo de revistas más productivas, sino para todas las fuentes registradas. Luego, los resultados allí publicados tienden a ser citados de manera más inmediata. Ya que el índice  $g$  busca compensar el volumen de citas puntual con el global,<sup>(32)</sup> también puede verse que *Human Molecular Genetics* logra el mejor balance con  $g = 105$ .

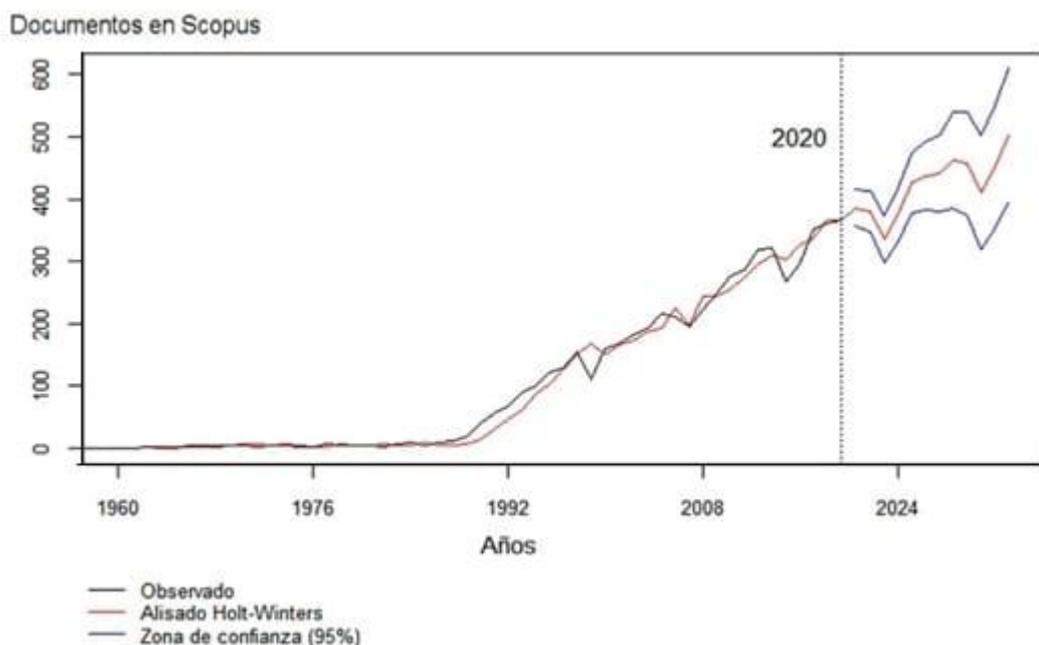
La extensión de los documentos publicados oscila entre una y 76 páginas (promedio = 8,21, desviación típica = 5,54). En 758 registros (13,40 %) está ausente al menos uno de los metadatos sobre el paginado, o bien aparecen errores en la entrada de la información. Las dificultades más frecuentes consisten en la inversión o supresión de dígitos, así como el intercambio u omisión de los números inicial y final. En algunos casos, la enumeración sigue un formato mixto que incluye caracteres no numéricos (C, N, E, S, e, T, RA, x, fundamentalmente). Si bien el formato CSV conserva estos datos, algunos paquetes computacionales pueden introducir errores.

### Ritmo de crecimiento en la producción científica

En el transcurso de 60 años la producción científica experimenta un progreso evolutivo ascendente, con una inflexión tangible a partir de 1993. Desde el año 1970 el modelo exponencial se ajusta a la función  $y = 11,93\exp(0,07x)$  ( $R^2 = 0,94$ ), con un porcentaje de variabilidad residual del 3,19 %. Este hecho se corresponde con la ley de Price, quien avizora que la ciencia avanza exponencialmente, desde una “pequeña ciencia” desarrollada por un grupo reducido de científicos, hacia una “gran ciencia”, regularmente financiada, institucionalizada y con carácter más universal.<sup>(33)</sup> La figura 1 ilustra los resultados de implementar la técnica de alisado de Holt-Winters, para un pronóstico a corto plazo de la serie temporal (coeficientes de suavizado: nivel  $\alpha = 0,42$ ; tendencia  $\beta = 0,21$  y estacionalidad  $\gamma = 1$ ).

Durante el período enmarcado entre 1993 y 2020, el crecimiento diacrónico concuerda incluso con un modelo lineal ( $y = 12,14x - 366,75$ ;  $R^2 = 0,97$  y alcanza mejor variabilidad residual del 0,11 %). Este desarrollo acelerado se corresponde con la identificación

gradual de las mutaciones causantes de las SCAs. A partir de 1991 se comienzan a descubrir los diferentes subtipos moleculares de ataxia (SCA 2, 3, 6, etc.), producidos por un incremento en el número de repeticiones del trinucleótido CAG en la región codificadora de los respectivos genes, conocidas como enfermedades poliglutamínicas. Dicho proceso abrió nuevas posibilidades en las investigaciones clínicas, moleculares, fisiopatológicas y terapéuticas.<sup>(1)</sup>



Fuente: Elaboración propia con R.

**Fig. 1** – Crecimiento diacrónico de la productividad científica.

También se experimentan descensos estacionarios tras intervalos de aproximadamente 7 años, con declives puntuales en 2002, 2009 y 2016. Probablemente, cada declive constituye el efecto de una desaceleración en el empleo de resultados proveniente de estudios genéticos en el período correspondiente. Ello sugiere complementar estas observaciones con análisis de contenido, lo cual constituye un terreno fértil para investigaciones posteriores. Con mirada al futuro, el

pronóstico sugiere sostenibilidad en el crecimiento global, que podría verse afectado por declives estacionales en 2023 (alrededor de 336 documentos, con intervalo de confianza [373; 298];  $p < 0,005$ ) y en 2030 (análogamente 412; [503; 320];  $p < 0,05$ ). Además, la etapa final del presente estudio coincide con el inicio de la pandemia de la COVID-19, lo cual también podría influir consustancialmente en el comportamiento de los volúmenes y contenidos de las investigaciones en SCAs.<sup>(3,34)</sup>

En estudios previos se ha observado que el volumen acumulativo de documentos en base de datos guarda un vínculo estrecho con las especificidades del saber científico.<sup>(35)</sup> Para las Ciencias Sociales y Humanísticas resulta efectivo el modelo de crecimiento doble exponencial de Gompertz, mientras que los campos relacionados con ciencia y tecnología se ajustan mejor al modelo de potencia. A tono con este hallazgo, se observa que el volumen acumulativo de documentos sobre SCAs en Scopus concuerda bien con una ley de potencia de la forma

$$y = a(x - 1960)^b$$

donde  $x \geq 1961$  representa el año,  $y$  la cantidad de documentos correspondientes, mientras que  $a$  y  $b$  son dos parámetros a estimar. Los resultados de la estimación producen un ajuste favorable ( $R^2 = 0,996$ ) y se reflejan en la tabla 3.

**Tabla 3** - Estimación de los parámetros del modelo de potencia

Parámetro	Estimación	Error estándar	Límite inferior*	Límite superior*
$a$	$17,12 \times 10^{-6}$	$5,00 \times 10^{-6}$	$72,92 \times 10^{-6}$	$26,96 \times 10^{-6}$
$b$	4,80	0,07	4,66	4,94

Leyenda: \* $p < 0,05$

Fuente: Elaboración propia.

Si bien el ajuste comporta una variabilidad residual de apenas el 0,2 %, el modelo no establece un resultado definitivo. El uso de extrapolaciones de tendencias para conjuntos de datos bibliométricos presenta desafíos, principalmente relacionados con

el tamaño de la muestra, la fiabilidad de la información y el esclarecimiento de factores subyacentes que influyen en el poder explicativo como fuerzas causales. Por este motivo, al emplear un modelo, es sugerente seguir la regla de no extender los pronósticos más allá del período de tiempo que comprenden los datos.<sup>(36)</sup> Además, si bien los valores  $R^2$  de los modelos exponencial y lineal resultan ligeramente inferiores al modelo de potencia, su empleo resulta más y de uso extendido.

### Paternalidad de la obra e impacto por índice de citas

Para el presente estudio, el análisis de la paternidad de un documento se basa en las interrelaciones nación-organización-investigador, mientras que el nivel de impacto se apoya en el índice de citas. Estados Unidos, Japón y Alemania tienen participación en no menos del 10 % de los documentos publicados, como muestra la tabla 4, contentiva del *ranking* para los primeros 20 países. Tomando en consideración el aspecto demográfico, países como Portugal, Países Bajos y Cuba presentan tasas de más de un documento por cada 100 mil habitantes. En general, la cantidad de documentos no solo responde al desarrollo económico del país, sino también al nivel de prevalencia de esta enfermedad.

**Tabla 4** - Ranking por países basado en la firma de documentos

N.º	País	Frecuencia*	Porcentaje*	Documentos con autor correspondiente	MCP Ratio <sup>†</sup>	Tasa <sup>‡</sup>
1	Estados Unidos	1693	29,94	1009	0,18	0,51
2	Japón	758	13,41	563	0,09	0,60
3	Alemania	614	10,86	326	0,40	0,73
4	Reino Unido	473	8,37	270	0,44	0,70
5	Italia	442	7,82	254	0,22	0,73
6	China	395	6,99	327	0,14	0,03
7	Francia	347	6,14	156	0,38	0,53
8	Brasil	282	4,99	221	0,33	0,13
9	Países Bajos	249	4,40	110	0,37	1,45
10	Canadá	247	4,37	102	0,32	0,65
11	España	217	3,84	103	0,18	0,46

12	India	170	3,01	125	0,16	0,01
13	Taiwán	165	2,92	128	0,12	0,07
14	Portugal	165	2,92	92	0,59	1,62
15	Cuba	137	2,42	81	0,58	1,21
16	Australia	125	2,21	72	0,28	0,49
17	Bélgica	112	1,98	40	0,35	0,97
18	Corea del Sur	96	1,70	81	0,12	0,19
19	México	84	1,49	43	0,44	0,07
20	Suiza	76	1,34	40	0,45	0,88

Leyenda: \*Los valores no son excluyentes, a causa de las coautorías entre investigadores de diferentes países; †El cociente MCP (*Multiple Country Publications*) refiere el índice de colaboración internacional; ‡Tasa de documentos por cada 100 mil habitantes, con base en la *World Population Prospects: The 2019 Revision* (United Nations, Department of Economic and Social Affairs).

*Fuente:* Elaboración propia.

Aparecen representadas 16841 organizaciones (instituciones, fundaciones y asociaciones, principalmente, con 130 registros vacíos). Nueve instituciones suman al menos un centenar de documentos: *University of Minnesota* (235), *INSERM - Institut National de la Santé et de la Recherche Médicale* (208), *Baylor College of Medicine* (189), *Universität Tübingen* (156), *Hôpitaux Universitaires Pitié Salpêtrière* (150), *CNRS - Centre National de la Recherche Scientifique* (122), *CIRAH - Centro para la Investigación y Rehabilitación de las Ataxias Hereditarias* (122), *Xiangya Hospital* (112) y *TMDU - Tokyo Medical and Dental University* (100). Con al menos una decena de documentos aparecen 39 organizaciones. En particular, el CIRAH es una institución multidisciplinaria cubana fundada en el año 2000 y pionera de su tipo en América Latina, diseñada para la atención integral a los pacientes que padecen esta enfermedad y para el desarrollo de investigaciones científicas afines.

En toda la muestra la cantidad de autores por documento oscila dentro del rango de 1 a 103 (moda = 2; mediana = 6; media = 6,61; desviación típica = 5,24). Hasta 1995 el *test* de Shapiro-Wilk no permite rechazar la hipótesis de normalidad en el número de coautores por documento en cada año ( $gI \leq 57$ ;  $0,73 \leq W \leq 0,94$ ;  $p > 0,05$ ) con excepción de 1980, 1989, 1991, 1993-1995 con pequeñas muestras. A partir de 1996, la prueba de bondad de ajuste de Kolmogorov-Smirnov conduce a rechazar la hipótesis de normalidad en todos los casos ( $gI \geq 68$ ;  $0,11 \leq Z \leq 0,20$ ;  $p < 0,01$ ), con muestras

superiores a 50. Este hecho sugiere un cambio en los hábitos investigativos reflejados en el reconocimiento de coautoría.

La figura 2 contiene los diagramas de cajas del número de coautores por cada año. Poco más de la mitad de los documentos (50,07 %) tienen entre tres y nueve autores y la amplia mayoría (95,53 %) agrupan a lo sumo 15 coautores. Puede observarse el incremento paulatino de valores atípicos, principalmente durante la última década (*outlier* máximo = 103). De modo general, se aprecia una tendencia paulatina hacia el incremento del trabajo colectivo e interdisciplinario, lo cual explica los resultados del test de Kolmogorov-Smirnov. Un hecho similar se ha constatado en otros estudios bibliométricos que reconocen el incremento paulatino del número de integrantes en los grupos de investigación, aunque no siempre el resultado reportado guarda relación estrecha con la complejidad de la investigación.<sup>(37,38,39)</sup>

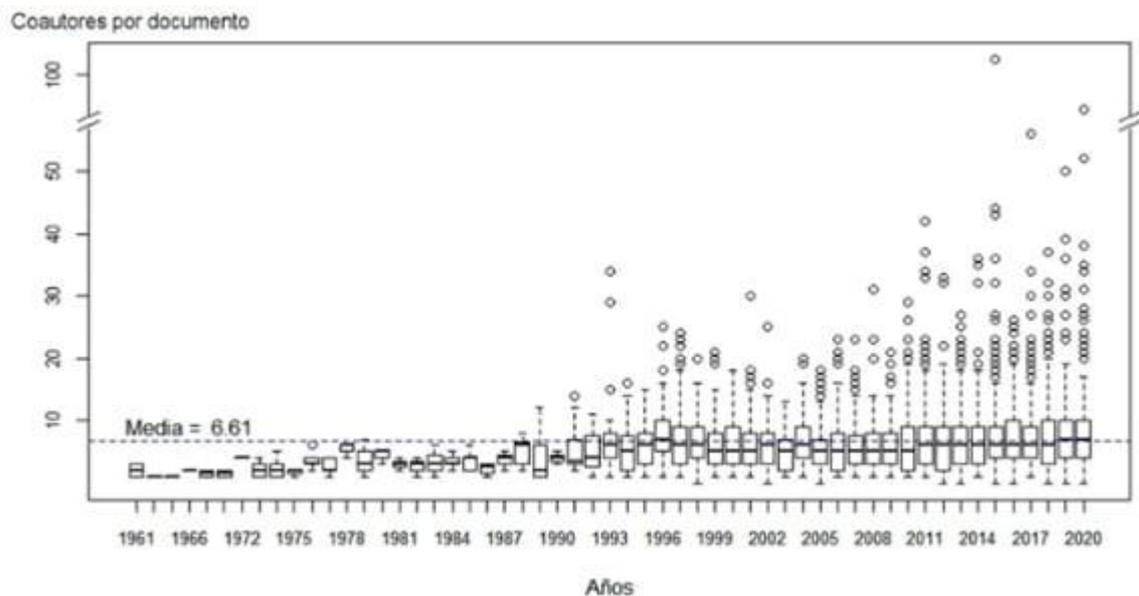


Fig. 2 – Cantidad de coautores por año.

Fuente: Elaboración propia con R.

Después de excluir 22 registros con datos perdidos se contabilizan 18271 autores en total. El rango de documentos por autor comprende entre uno y 132 trabajos (moda = mediana = 1, media = 2,05, desviación típica = 4,57). Los autores con presencia en un

único documento totalizan 13 241 (72,47 %), mientras que 17 054 (93,34 %) firman menos de cinco trabajos. Este hecho es consecuente con la ley de Lotka, relacionada con la distribución de autores según su productividad.<sup>(40)</sup> Un análisis de regresión produce el valor  $A = 13238,05$  (error típico = 7,66; intervalo de confianza = [13222,79; 13253,31];  $p < 0,05$ ) y  $C = 2,44$  (error típico = 0,0036; intervalo de confianza = [2,43; 2,44];  $p < 0,05$ ) para el modelo generalizado  $A_n = A/(n^C)$ , donde  $A_n$  es el número de autores con  $n$  firmas.<sup>(41)</sup> El modelo teórico se ajusta razonablemente a las observaciones (suma ajustada de cuadrados = 177900190,34;  $gl = 73$ ;  $R^2 > 0,99$ ). Por ejemplo,  $A_{10} = 13238,05/(10^{2,44}) \approx 48,64$  y el valor observado es 51 (error típico = 4,84 %). La tabla 5 muestra la primera decena de autores más prolíficos, junto al impacto de sus obras por índice de citas dentro de Scopus.

**Tabla 5 - Ranking de autores más productivos e impacto por índices de citas en Scopus**

N.º	Autor e ID	Afilación*	Indicadores locales <sup>†</sup>						Indicadores globales <sup>‡</sup>		
			Total	Citas	Año inicial	Índice <i>h</i>	Índice <i>m</i>	Índice <i>g</i>	Total	Citas	Índice <i>h</i>
1	Bricé Alexis, 56401594400	Hôpitaux Universitaires Pitié Salpêtrière	132	8275	1993	58	2,07	98	879	72407	126
2	Orr Harry T., 7102615743	University of Minnesota	122	14635	1987	58	1,71	119	260	25927	84
3	Zoghbi Huda Yahya, 7102095308	Baylor College of Medicine	119	16504	1988	63	1,91	117	416	53976	116
4	Dürr Alexandra, 24741153500	Hôpitaux Universitaires Pitié Salpêtrière	118	5680	1993	28	1,00	54	545	42039	103
5	Schöls Lüdger, 7005610155	Hertie-Institut für klinische Hirnforschung	115	5750	1995	46	1,77	87	454	21903	74
6	Klockgether Thomas, 26643063400	Universität Bonn	115	5087	1993	50	1,79	85	518	29170	91
7	Ashizawa Tetsuo, 7103079513	Weill Cornell Medical College	109	3799	1997	34	1,42	65	290	14137	59

8	Stévanin Giovanni, 56385311700	École Pratique des Hautes Études	97	6231	1993	45	1,61	86	280	15266	66
9	Pulst Stefan Matthias, 7004901399	University of Utah School of Medicine	92	2478	1993	37	1,32	73	276	11947	56
10	Velázquez-Pérez Luis, 6602512664	Academia de Ciencias de Cuba, CIRAH	91	917	2007	22	1,57	15	123	1909	24

Leyenda: \*Se refiere a la afiliación consignada en el documento más reciente; †Basados en SJR (*Scimago Journal & Country Rank*), respecto a la base de datos Scopus; ‡Basados en el análisis intrínseco de la matriz de datos BibTeX en el campo de las SCAs.

Fuente: Elaboración propia.

El volumen total de citas asciende a 155990 (moda = 1; mediana = 12; media = 31,53; desviación típica = 66,92). Existen 4948 obras con al menos una cita (87,50 %) y de los 706 documentos no citados, un total de 433 (61,33 %) son anteriores al último quinquenio. Respecto a los índices  $i10^k$ , en el sentido de Teixeira da Silva,<sup>(42)</sup> un monto de 2872 registros cuenta con al menos diez citas (índice  $i10$ ; 50,79 %); 306 sobrepasan un centenar de citas (índice  $i100$ ; 5,41 %) y tres superan las mil citas (índice  $i1000$ ; 0,05 %). Esto significa que el corpus documental sobre este campo del saber científico ha alcanzado madurez, durante seis décadas de registros en la base de datos Scopus. En particular, el trabajo más citado se titula *Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1* y fue publicado en la revista *Nature Genetics*.<sup>(2)</sup>

Respecto al volumen acumulado de citas a revistas, el análisis de regresión produce mejor ajuste bajo el modelo log-normal ( $R^2 = 0,9994$ ) que bajo el supuesto clásico log-lineal ( $R^2 = 0,6473$ ). Este hallazgo es coherente con un resultado de Ravichandra Rao, quien arriba a la misma conclusión a partir de un estudio metaanalítico basado en 12 investigaciones previas.<sup>(43)</sup> Para el caso de las investigaciones sobre SCAs, se ordenan las revistas conforme al acumulado de citas de todos sus documentos. Así resulta la serie  $x_1 < x_2 < \dots < x_{968}$ , donde se excluyen 214 fuentes carentes de cita ( $214 = 1182 - 968$ ) y en el caso donde las fuentes  $x_i$  y  $x_j$  tienen igual cantidad de citas acumuladas, entonces se disponen acorde al orden lexicográfico. A partir de aquí, si  $F(x_i)$  es la cantidad de citas acumuladas hasta la fuente  $x_i$ , entonces  $F(x_i) \sim 4,06 \times 10^{(-7)} M[\ln(x_i)]$ ;

$\mu = 6,96; \sigma = 4,44] + 0,60$ . Por ejemplo, el acumulado esperado para las diez fuentes más citadas es  $F(10) \approx 63034$ , mientras que el valor real es exactamente 58857, con un error relativo del 7,10 %. Análogamente,  $F(20) \approx 83918$  y el valor observado 84161 difiere apenas en el 0,29 %.

Después de estratificar la secuencia de fuentes ordenadas en tres grupos de igual número de referencias, se obtiene la proporción 9:31:928. Esta última no satisface la forma geométrica clásica  $1 : n : n^2$ , pero ilustra la ley de dispersión Bradford, al demarcar la existencia de un núcleo de fuentes bibliográficas con mayor impacto.<sup>(44)</sup> Diez de ellas comprenden el 37,73 % de las citas y 17 sobrepasan la mitad de todas las referencias para este campo de investigación en Scopus. Las diez revistas más citadas son: *Nature Genetics* (11794), *Human Molecular Genetics* (10743), *Neurology* (6190), *Brain* (5703), *Annals of Neurology* (4705), *Proceedings of the National Academy of Sciences of the United States of America* (4164), *Journal of Neuroscience* (4005), *Nature* (3870), *American Journal of Human Genetics* (3843) y *The Cerebellum* (3840).

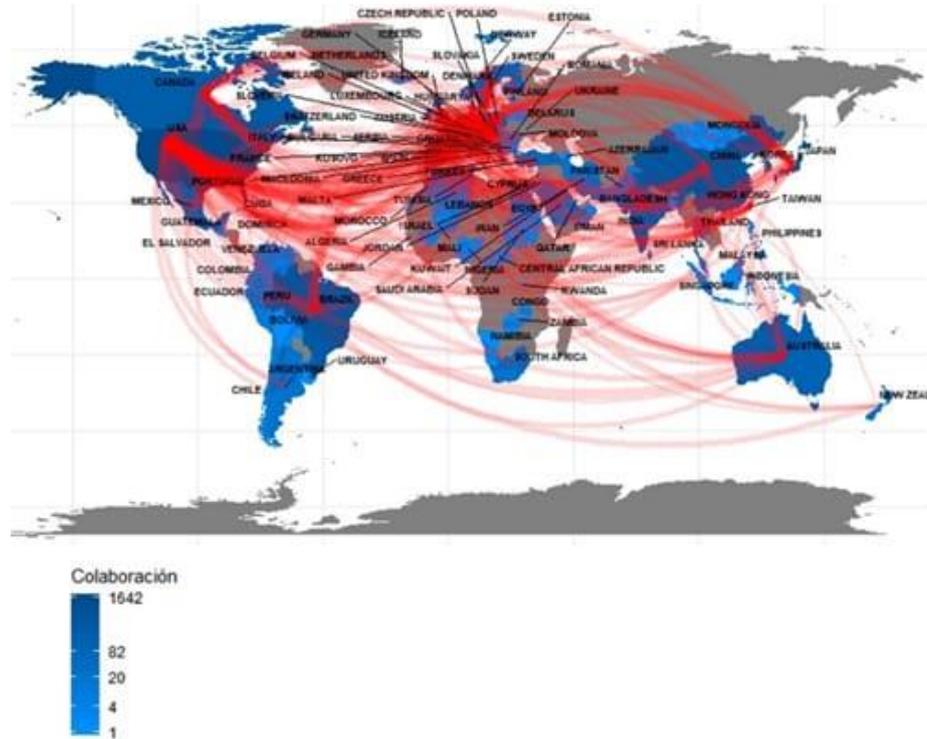
La selección y empleo de referencias sirve de soporte para sostener los argumentos científicos, enmarcar el planteamiento de un problema, revelar la novedad de los aportes, entre otros elementos de la investigación.<sup>(45,46)</sup> En el presente estudio, si bien existen 393 registros carentes de bibliografía, ello apenas constituye un 6,95 % del total, sin rebasar el 10 % de los registros por año desde 2002. En algunos casos se trata de obras que ciertamente no incluyen citas, pero un muestreo adicional de los artículos referenciados revela que mayormente consiste en información ausente. Por tal motivo, estos registros no se incluyen en el análisis. Los 5261 documentos considerados tienen un rango de referencias entre 1 y 265, donde el 50 % contiene entre 26 y 84. Después de calcular cada razón entre el número de documentos y el volumen de referencias del año correspondiente, se observa un comportamiento lineal ascendente ( $y = 0,96x - 5,29; R^2 = 0,87$ ), lo cual indica que existe tendencia sostenida hacia el incremento del número promedio de referencias en cada investigación reportada con el paso del tiempo.

La consideración de las referencias como un todo e independientemente del año de aparición produce una muestra con cierto comportamiento asimétrico positivo ( $\tilde{\mu}_3 = 1,86$ ) y leptocúrtico ( $\kappa = 6,91$ ). El test de bondad de ajuste de Kolmogorov-Smirnov

conduce a rechazar la hipótesis de normalidad ( $D = 0,98$ ;  $p < 0,01$ ). En particular se aprecia discordancia respecto a la distribución log-normal ( $D = 0,07$ ;  $p < 0,01$ ), incluso con peor ajuste que otras distribuciones tales como Gamma, Weibull, ji al cuadrado, exponencial, logística y uniforme. Este comportamiento no concuerda con los hallazgos de un estudio desarrollado por Brzezinski, donde log-normal constituyó uno de los mejores modelos para distribuciones de citas.<sup>(47)</sup>

### **Redes de colaboración y estructura del flujo del conocimiento**

La colaboración científica tiene una expresión concreta en los índices de coautoría. La figura 3 refleja el complejo entramado mundial de cooperación, en investigaciones relacionadas con las SCAs. La visualización se basa en la existencia de al menos un par de documentos en colaboración internacional, con un rango que alcanza el valor de 1642 para el caso de Estados Unidos. En la escala de intensidad del color azul puede verse que los tres primeros cuartiles comprenden hasta 82 documentos en la amplia mayoría de los países. La productividad se extiende principalmente por regiones como Europa Occidental y Meridional, las Américas, Asia Oriental, Sudoriental y Meridional, Australia y Nueva Zelanda, así como África Septentrional y Meridional. Sin embargo, las mayores redes de colaboración conectan a Europa con Estados Unidos, Canadá, Brasil y Japón, fundamentalmente.



Fuente: Elaboración propia con R.

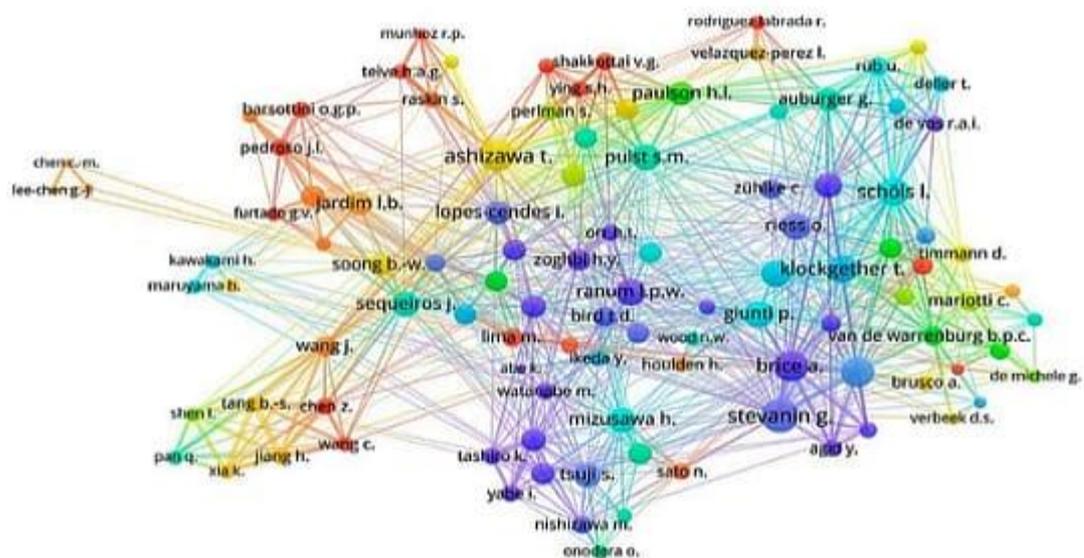
**Fig. 3** – Redes de colaboración entre países.

El análisis de la red de coautorías aporta un conjunto de indicadores que describen su estructura, conforme a la nacionalidad. Por ejemplo, el tamaño  $n = 92$  indica que casi todos los países tienen presencia en el proceso de cooperación (95,83 %). El diámetro  $d = 4$  revela que la distancia geodésica más larga (longitud del camino más corto entre dos nodos) provee a la red de una cercanía relativa respecto a su centro, ubicado aproximadamente en los Estados Unidos (grado de centralización  $C = 0,48$ , longitud promedio de ruta  $l = 2,01$ ). Sin embargo, la densidad es relativamente baja ( $D = 0,18$ ), así que la red tiende a ser estrellada y poco conexa. Por tanto, predomina la cooperación aislada de muchos países respecto a unos pocos.

Con base en la cooperación entre colegas, la red de coautorías refleja la consolidación de grupos de investigación relativamente estables en el tiempo. El número de nodos  $N = 17754$  (cantidad de autores) dota la red de un elevado nivel de complejidad, con diámetro  $D = 12$ , que sugiere la existencia de distancias considerables entre algunos componentes. Sin embargo, la longitud promedio de una ruta ( $l = 4,23$ ) ocupa cerca de

la tercera parte del diámetro. La densidad ( $D = 0,001$ ) y el grado de centralización ( $C = 0,04$ ) son notablemente bajos, lo cual indica la ausencia de un liderazgo o preeminencia de algún subgrupo, así como conexión más marcada en los subgrupos, en detrimento de una cohesión global. De forma general, se contabilizan 23 clústeres que agrupan entre dos y 114 autores.

Para clarificar la esencia estructural de esta compleja red de coautorías es conveniente seleccionar un subconjunto representativo en el sentido de su productividad e impacto científico. La figura 4 muestra una parte significativa de la red, donde han visualizado 106 autores con al menos 20 documentos, 20 vínculos fuertes y 200 citaciones. El procesamiento con VOSviewer produce un total de 11 clústeres que reflejan nexos de colaboración ( $21 + 13 + 13 + 12 + 9 + 8 + 7 + 7 + 6 + 6 + 5 = 106$  autores; método de normalización = modularidad *LinLog*, atracción = 2, repulsión = -1, peso en la escala de visualización = fuerza total del enlace). El grosor de las líneas es proporcional a la intensidad del trabajo conjunto, mientras que el espectro de colores refleja las relaciones cronológicas de cooperación durante la década 2005-2015. Los grupos representados con colores fríos son más antiguos, mientras que los más recientes se muestran con colores cálidos.



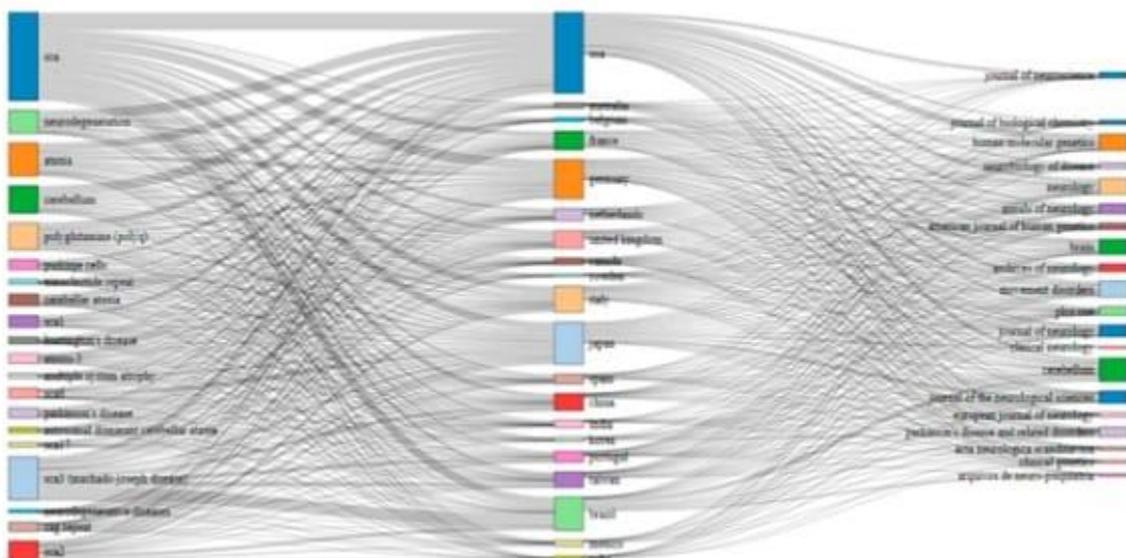
Fuente: Elaboración propia con VOSviewer.

Fig. 4 – Redes de colaboración entre autores.

Aunque el promedio de autores por documentos es ligeramente superior a seis en todos los documentos compilados, este indicador no refleja con claridad las relaciones más estables y consolidadas de trabajo. La figura 4 revela dos cuestiones que pueden ayudar a explicar este proceso con mayor claridad. En efecto, las relaciones más sólidas se aprecian entre dúos de investigadores, mientras que el predominio de triángulos refleja que de los seis autores promedio en un mismo trabajo, cerca de la mitad coinciden en otras publicaciones.

Los contenidos fundamentales de las investigaciones publicadas pueden analizarse con ayuda de los metadatos correspondientes a los títulos, resúmenes y palabras clave. En el presente estudio no se encontraron diferencias sustanciales entre la información proporcionada por las palabras clave (de autor e indizadas) y los restantes metadatos mencionados. La figura 5 consiste en un diagrama de Sankey que refleja un flujo de producción y socialización del conocimiento. En la columna izquierda aparecen las 20 palabras claves más recurrentes, relacionadas con investigaciones sobre SCAs. La columna central contiene los 20 países donde fundamentalmente se originan estas investigaciones. Finalmente, en la columna derecha se muestran las 20 revistas donde se concentran las publicaciones de los resultados correspondientes.

Nueve de las diez revistas más productivas aparecen en esta relación (ver tabla 2), donde la mayor cantidad de coincidencias se aprecia en las revistas *The Cerebellum*, *Movement Disorders*, *Human Molecular Genetics*, *Neurology*, y *Brain*. De forma ligeramente alejada del núcleo de fuentes, aparece la revista *Journal of Neuroscience*, de la Sociedad de Neurociencia. En general, estas revistas abordan temáticas específicas relacionadas con la enfermedad. Por ejemplo, *The Cerebellum* se especializa en la estructura que más se daña en las ataxias, mientras que *Movement Disorders* constituye el órgano oficial de la Sociedad Internacional de Parkinson y Trastornos del Movimiento, y justo en este último grupo de enfermedades neurológicas se engloban las SCAs.



Fuente: Elaboración propia con VOSviewer.

Fig. 5 – Origen y fuentes documentales de los aspectos más investigados.

La información provista por la figura 5, junto a las frecuencias observadas en las palabras clave permiten concluir que los tipos más frecuentes en las investigaciones registradas son SCA1, SCA2, SCA3 y SCA6, precisamente las ataxias de mayor prevalencia, lo cual concuerda con las conclusiones de Diallo y otros en un estudio metaanalítico.<sup>(1)</sup> Junto a estos tipos, el SCA17 (también conocido como HDL4: *Huntington-like Disease type 4*) constituye un tipo que se asemeja a la enfermedad de Huntington, el cual es menos frecuente pero más ampliamente estudiado.

## Conclusiones

Las SCAs constituyen enfermedades neurodegenerativas cuyos efectos implican un deterioro sensible en la salud y, por consiguiente, en la calidad de vida de los pacientes. Con la presente investigación se explora la producción científica relacionada con dicha afectación, a partir de una recopilación de naturaleza censal-documental y con fuente en Scopus. El conjunto de metadatos recuperados corresponde a 5654 documentos y comprende las últimas seis décadas. Por tanto, el análisis proporciona información basada en un volumen considerable de evidencias empíricas, lo cual proporciona al estudio un nivel razonable de fiabilidad.

Aspectos tales como la ausencia de palabras clave, deficiencias de normalización y omisión de metadatos, constituyen variables que han requerido de control sistemático, a fin de disminuir su incidencia negativa en la validez de los resultados. No obstante, existen elementos que proveen al estudio de consistencia, en el sentido de constatar principios bibliométricos tales como las leyes de Price, de Lotka y de Bradford y también de concordar con hallazgos descritos por otros autores. Por ejemplo, pudo constatar que el modelo de regresión que más se ajusta a la distribución de citas a revistas específicas y también a la distribución de referencias por documento, es el modelo *log-normal*.<sup>(43,47)</sup> Asimismo, el modelo de potencia produce una mejor regresión para el crecimiento diacrónico anual, aunque ello no supera los beneficios prácticos de los modelos lineal y exponencial.<sup>(35)</sup>

El corpus documental sobre las SCAs refleja la consolidación de grupos de investigación relativamente estables, encabezados por líderes científicos y caracterizados por la ampliación de la colaboración internacional y por el trabajo colectivo e interdisciplinario. Se aprecia una tendencia al aumento del número de referencias en cada documento, agrupadas generalmente dentro de la corriente principal. Los mayores volúmenes productivos se concentran en países desarrollados, junto a países en vías de desarrollo donde existen elevados niveles de prevalencia en esta enfermedad. Los tipos de ataxia más investigados coinciden con aquellos de mayor prevalencia a escala mundial.

Algunas observaciones demarcan nuevos caminos para investigaciones ulteriores. Por ejemplo, es importante esclarecer los vínculos entre el crecimiento diacrónico estacional y los estudios genéticos que podrían haber influido en los procesos de aceleración/desaceleración local, dentro del desarrollo exponencial global. También es trascendente ahondar en los campos de investigación específicos de cada clúster identificado, pues ello ayudaría a comprender con más claridad las tendencias de investigación y su evolución con el paso del tiempo. En general, estos aspectos responden a un complemento consustancial para los estudios bibliométricos, relacionado con el análisis de contenido.

## Disponibilidad de los datos

Ambas bases de datos, en formatos CSV y BIB aparecen disponibles en *Figshare*, junto a la totalidad de los códigos escritos sobre lenguaje R para la presente investigación.<sup>(48,49)</sup>

## Referencias bibliográficas

1. Diallo A, Jacobi H, Tezenas du Montcel S, Klockgether T. Natural history of most common spinocerebellar ataxia: a systematic review and meta-analysis. *Journal of Neurology*. 2021;268(8):2749-56. DOI: <https://doi.org/10.1007/s00415-020-09815-2>
2. Orr HT, Chung M-Y, Banfi S, Kwiatkowski TJJr, Servadio A, Beaudet AL, *et al*. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nature Genetics*. 1993;4(3):221-6. DOI: <https://doi.org/10.1038/ng0793-221>
3. Velázquez L, Vázquez Y, Rodríguez R. Ataxias hereditarias y COVID-19: posibles implicaciones fisiopatológicas y recomendaciones. *Anales de la Academia de Ciencias de Cuba*. 2020 [acceso 08/12/2023];10(2). Disponible en: <http://www.revistaccuba.cu/index.php/revacc/rt/prINTERfriendly/801/830>
4. Velázquez L, Rodríguez R. Riluzole and spinocerebellar ataxia type 2: the ATRIL trial. *The Lancet Neurology* (preprint). 2022;21(3):e3. DOI: [https://doi.org/10.1016/S1474-4422\(22\)00028-X](https://doi.org/10.1016/S1474-4422(22)00028-X)
5. Teive HAG, Meira AT, Camargo CE, Munhoz RP. The geographic diversity of spinocerebellar ataxias (SCAs) in the Americas: a systematic review. *Movement Disorders Clinical Practice*. 2019;6(7):531-40. DOI: <https://dx.doi.org/10.1002%2Fmdc3.12822>
6. Ortega G, Abenza MJ, Serrano C, Axpe R, Arpa FJ, Adarmes AD, *et al*. Mapa epidemiológico transversal de las ataxias y paraparesias espásticas hereditarias en España. *Neurología* (preprint). 2021. DOI: <https://doi.org/10.1016/j.nrl.2021.01.006>

7. Martínez MC, Anglada LC, Cruz M. La atención educativa integral a escolares con ataxia en contextos inclusivos. *Roca*. 2021 [acceso 08/12/2023];5(1):36-48. Disponible en: <https://revistas.udg.co.cu/index.php/redel/article/view/2155/4095>
8. Lacorte E, Bellomo G, Nuovo S, Corbo M, Vanacore N, Piscopo P. The use of new mobile and gaming technologies for the assessment and rehabilitation of people with ataxia: a systematic review and meta-analysis. *The Cerebellum*. 2021;20(3):361-73. DOI: <https://doi.org/10.1007/s12311-020-01210-x>
9. Novis LE, Spitz M, Jardim M, Raskin S, Teive HAG. Evidence and practices of the use of next generation sequencing in patients with undiagnosed autosomal dominant cerebellar ataxias: a review. *Arquivos de Neuropsiquiatria*. 2020 [acceso 08/12/2023];78(9):576-85. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/32725052/>
10. Ruano L, Melo C, Silva MC, Coutinho P. The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies. *Neuroepidemiology*. 2014;42(3):174-83. DOI: <https://doi.org/10.1159/000358801>
11. Serrano-López AE, Martín-Moreno C, Sanz-Casado E. Análisis bibliométrico de la influencia de la Genética en enfermedades raras a partir de las bases de datos Pubmed y SCI (2000-2009). *Revista Española de Documentación Científica*. 2013;36(4):e024. DOI: <https://doi.org/10.3989/redc.2013.4.970>
12. Shukla R. Indian research output on genetic disorder publication using the Scopus database: A scientometric study. *COLLNET Journal of Scientometrics and Information Management*. 2019;13(1):91-102. DOI: <https://doi.org/10.1080/09737766.2018.1550044>
13. Karthikeyan G, Manoharan A, Swaminathan S. A scientometric study on neuro science with special reference to growth of literature. *Indian Journal of Information Sources and Services*. 2019 [acceso 08/12/2023];9(1):77-9. Disponible en: <http://www.trp.org.in/wp-content/uploads/2019/05/MANLIBNET-IJISS-Vol.9-No.S1-February-2019-pp.-77-79.pdf>
14. Liao Z, Wei W, Yang M, Kuang X, Shi J. Academic publication of neurodegenerative diseases from a bibliographic perspective: a comparative scientometric analysis.

Frontiers in Aging Neuroscience. 2021;13:722944. DOI:  
<https://doi.org/10.3389/fnagi.2021.722944>

15. Pajo AT, Espiritu AI, Jamora RDG. Scientific impact of movement disorders research from Southeast Asia: A bibliometric analysis. *Parkinsonism & Related Disorders*. 2020;81:205-212. DOI: <https://doi.org/10.1016/j.parkreldis.2020.10.043>

16. Tesio L, Gamba C, Capelli A, Franchignoni FP. Rehabilitation: The Cinderella of neurological research? A bibliometric study. *Italian Journal of Neurological Sciences*. 1995;16(7):473-7. DOI: <https://doi.org/10.1007/bf02229325>

17. Fan J, Gao Y, Zhao N, Dai R, Zhang H, Feng X, *et al*. Bibliometric analysis on COVID-19: A comparison of research between English and Chinese studies. *Frontiers in Public Health*. 2020;8:477. DOI: <https://doi.org/10.3389/fpubh.2020.00477>

18. Shamszadeh S, Asgary S, Nosrat A. Regenerative endodontics: A scientometric and bibliometric analysis. *Journal of Endodontics*. 2019;45(3):272-80. DOI: <https://doi.org/10.1016/j.joen.2018.11.010>

19. López-Illescas C, de Moya-Anegón F, Moed HF. Coverage and citation impact of oncological journals in the Web of Science and Scopus. *Journal of Informetrics*. 2008;2(4):304-16. DOI: <https://doi.org/10.1016/j.joi.2008.08.001>

20. Martín-Martín A, Orduna-Malea E, Delgado E. Coverage of highly-cited documents in Google Scholar Web of Science and Scopus: a multidisciplinary comparison. *Scientometrics*. 2018;116(3):2175-88. DOI: <https://doi.org/10.1007/s11192-018-2820-9>

21. Baas J, Schotten M, Plume A, Côté G, Karimi R. Scopus as a curated high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*. 2020;1(1):377-86. DOI: [https://doi.org/10.1162/qss\\_a\\_00019](https://doi.org/10.1162/qss_a_00019)

22. Schotten M, El Aisati M, Meester WJN, Steiginga S, Ross CA. A brief history of Scopus: The world's largest abstract and citation database of scientific literature. En: Cantu-Ortiz FJ, editor. *Research Analytics Boosting University Productivity and*

Competitiveness through Scientometrics. New York: Taylor & Francis Group LLC; 2018. p 31-58. DOI: <https://doi.org/10.1201/9781315155890>

23. Zhu J, Liu W. A tale of two databases: the use of Web of Science and Scopus in academic papers. *Scientometrics*. 2020;123(1):321-35. DOI: <https://doi.org/10.1007/s11192-020-03387-8>

24. The R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing [Computer Software]. Version 4.2.0. Vienna, Austria; 2022. Disponible en: <https://www.R-project.org/>

25. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*. 2017;11(4):959-75. DOI: <https://doi.org/10.1016/j.joi.2017.08.007>

26. Van Eck NJ, Waltman L. VOSviewer. Visualizing Scientific Landscapes [Computer Software]. Version 1.6.18. Leiden University: Centre for Science and Technology Studies; 2022. Disponible en: <https://www.vosviewer.com/>

27. León-Vargas F, Arango JA, Luna HJ. Two decades of research in artificial pancreas: Insights from a bibliometric analysis. *Journal of Diabetes Science and Technology*. 2022;16(2):434-45. DOI: <https://doi.org/10.1177/19322968211005500>

28. Boudry C, Baudouin C, Mouriaux F. International publication trends in dry eye disease research: A bibliometric analysis. *The Ocular Surface*. 2018;16(1):173-9. DOI: <https://doi.org/10.1016/j.jtos.2017.10.002>

29. Ruiz-Pomeda A, Álvarez-Peregrina C, Povedano-Montero F. Bibliometric study of scientific research on optometric visual therapy. *Journal of Optometry*. 2020;13(3):191-7. DOI: <https://doi.org/10.1016/j.optom.2019.12.007>

30. Kim E. A comparative analysis of research on LIS information behavior and health information seeking behavior. *Journal of the Korean BIBLIA Society for Library and Information Science*. 2019;30(2):167-87. DOI: <https://doi.org/10.14699/KBIBLIA.2019.30.2.167>

31. Hirsch JE. An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America–PNAS. 2005;102(46):16569-72. DOI: <https://doi.org/10.1073/pnas.0507655102>
32. Egghe L. Theory and practice of the g-index. Scientometrics. 2006;69(1):131-52. DOI: <https://doi.org/10.1007/s11192-006-0144-7>
33. Price DJS. Little Science Big Science. New York: Columbia University Press; 1963.
34. González-Garcés Y, Domínguez-Barríos Y, Zayas-Hernández A, Sigler-Villanueva AA, Canales-Ochoa N, Hernández-Oliver MO, *et al.* Impacts of the COVID-19 Pandemic on the Mental Health and Motor Deficits in Cuban Patients with Cerebellar Ataxias. The Cerebellum. 2021;20(6):896-903. DOI: <https://doi.org/10.1007/s12311-021-01260-9>
35. Egghe L, Ravichandra Rao IK. Classification of growth models based on growth rates and its applications. Scientometric. 1992;25(1):5-46. DOI: <https://doi.org/10.1007/BF02016845>
36. Suominen A, Seppänen M. Bibliometric data and actual development in technology life cycles: flaws in assumptions. Foresight. 2014;16(1):37-53. DOI: <https://doi.org/10.1108/FS-03-2013-0007>
37. Lutnick E, Cusano A, Sing D, Curry EJ, Li X. Authorship proliferation of research articles in top 10 orthopaedic journals: A 70-year analysis. Journal of the American Academy of Orthopaedic Surgeons: Global Research & Reviews. 2021;5(9):e2100098. DOI: <https://dx.doi.org/10.5435/JAAOSGlobal-D-21-00098>
38. Schrock JB, Kraeutler MJ, McCarty EC. Trends in authorship characteristics in The American Journal of Sports Medicine 1994 to 2014. The American Journal of Sports Medicine. 2016;44(7):1857-60. DOI: <https://doi.org/10.1177/0363546516639955>
39. Papatheodorou SI, Trikalinos TA, Ioannidis JPA. Inflated numbers of authors over time have not been just due to increasing research complexity. Journal of Clinical Epidemiology. 2008;61(6):546-51. DOI: <https://doi.org/10.1016/j.jclinepi.2007.07.017>
40. Lotka AJ. The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences. 1926 [acceso 10/02/2022];16(12):317-23. Disponible en: <https://www.jstor.org/stable/i24527553>

41. Kawamura M, Thomas CDL, Tsurumoto A, Sasahara H, Kawaguchi Y. Lotka's law and productivity index of authors in a scientific journal. *Journal of Oral Science*. 2000;42(2):75-8. DOI: <https://doi.org/10.2334/josnusd.42.75>
42. Teixeira da Silva JÁ. The i100-index i1000-index and i10000-index: expansion and fortification of the Google Scholar h-index for finer-scale citation descriptions and researcher classification. *Scientometrics*. 2021;126(4):3667-72. DOI: <https://doi.org/10.1007/s11192-020-03831-9>
43. Ravichandra Rao IK. An analysis of Bradford multipliers and a model to explain law of scattering. *Scientometrics*. 1998;41(1-2):93-100. DOI: <https://doi.org/10.1007/BF02457970>
44. Bradford SC. Sources of information on specific subjects. *Journal of Information Science*. 1985;10(4):176-80. DOI: <https://doi.org/10.1177/016555158501000407>
45. Bohannon R. How to find relevant references for a publication. *Psychotherapy Practice*. 1988;4(1):41-4. DOI: <https://doi.org/10.3109/09593988809161441>
46. Sim J. The use of bibliographic references. *Psychotherapy Practice*. 1989;5(1):25-32. DOI: <https://doi.org/10.3109/09593988909037758>
47. Brzezinski M. Power laws in citation distributions: evidence from Scopus. *Scientometrics*. 2015;103(1):213-28. DOI: <https://doi.org/10.1007/s11192-014-1524-z>
48. Cruz M, Rodríguez R, Velázquez LC. BibliometricAtaxias [Bases de datos]. Figshare. 2022. DOI: <https://doi.org/10.6084/m9.figshare.19092233>
49. Cruz M, Rodríguez R, Velázquez LC. StatisticsAtaxias [Códigos en lenguaje R]. Figshare. 2022. DOI: <https://doi.org/10.6084/m9.figshare.19092257>

### Conflicto de intereses

Los autores declaran que no tienen conflicto de intereses.

---

**Contribuciones de los autores**

*Conceptualización:* Miguel Cruz-Ramírez, Roberto Rodríguez-Labrada, Luis Clodovaldo Velázquez-Pérez.

*Curación de datos:* Miguel Cruz-Ramírez.

*Análisis formal:* Roberto Rodríguez-Labrada, Luis Clodovaldo Velázquez-Pérez.

*Investigación:* Miguel Cruz-Ramírez, Roberto Rodríguez-Labrada, Luis Clodovaldo Velázquez-Pérez.

*Metodología:* Miguel Cruz-Ramírez, Roberto Rodríguez-Labrada, Luis Clodovaldo Velázquez-Pérez.

*Software:* Miguel Cruz-Ramírez.

*Supervisión:* Luis Clodovaldo Velázquez-Pérez.

*Visualización:* Miguel Cruz-Ramírez.

*Redacción — borrador original:* Miguel Cruz-Ramírez, Roberto Rodríguez-Labrada, Luis Clodovaldo Velázquez-Pérez.

*Redacción — revisión y edición:* Miguel Cruz-Ramírez, Roberto Rodríguez-Labrada, Luis Clodovaldo Velázquez-Pérez.

**Financiación**

Este trabajo ha sido financiado en el marco del proyecto de investigación PN305LH013 del programa nacional de Neurociencias y Neurotecnología, titulado *Identificación de biomarcadores novedosos tempranos en enfermedades poliglutamínicas. Estudio longitudinal en una cohorte extensa.*