

El SIDA y su productividad científica en la base de datos SciELO entre 1997 - 2003: estudio bibliométrico

Por:

Richard Chiroque Solano y Ariela Medina Valdivia

Estudiantes de pregrado de la Escuela Académico Profesional de Bibliotecología. Facultad de Letras y Ciencias Humanas de la Universidad Nacional Mayor de San Marcos (Perú).

Correo electrónico: richardchiroque@lycos.es ; amedinavaldivia@yahoo.es

Resumen

Se analizó la producción científica de un conjunto de revistas sobre el SIDA, en la base de datos SciELO, durante los años 1997-2003, aplicando el modelo matemático de Bradford. La investigación estuvo centrada en la identificación de indicadores relacionados con productividad de las revistas, cobertura geográfica e idiomática, y producción anual. Se ubicó en total 315 registros. Los principales resultados fueron: a) 3 revistas formaron el núcleo de Bradford. b) Brasil y Chile fueron los países con mayor producción. c) El idioma portugués obtuvo un alto índice de incidencia. d) Los años 2001 y 2002 fueron los más productivos respectivamente.

Palabras clave: Bibliometría, SciELO, SIDA, Modelo matemático de Bradford

1.1. Formulación del problema

La producción científica en salud en Latinoamérica es relativamente poco expresiva en términos cuantitativos y altamente concentrados. América Latina es responsable apenas del 2% de artículos indizados en MEDLINE (6758 artículos para un total de 421.833 en 1998). En lo que se refiere a la distribución de esta producción, Argentina (1.168), Brasil (2882) y México (1.271), fueron responsables del 80% de las publicaciones de América Latina y el Caribe. Perú tuvo 429 artículos en 1998 y 436 en 1999.¹

Esta investigación en el área de las Ciencias de la Salud a nivel latinoamericano, aborda el tema del SIDA en la base de datos SciELO y abarca el período 1997-2003. El propósito es conocer si el conjunto de revistas a analizar se ajustan al modelo matemático de Bradford, e indagar la productividad científica y el estado de investigación en la región, así como indicadores que se desprenden de ello (idioma y cobertura geográfica). Para establecer el grado de investigación y el crecimiento de la producción científica referente al SIDA se realizaron búsquedas en diversas base de datos en línea y se obtuvo algunas referencias a estudios bibliométricos sobre SIDA, pero ningún aporte sobre la aplicación del modelo matemático de Bradford en este campo específico de las ciencias de la salud.

¹ OPS. Informe de salud de las Américas 2002

Entre ellos tenemos, *Análisis bibliométrico de la producción científica española sobre SIDA a través de la base de datos IME durante el período 1983-1991* (1993), en el que se identifica el crecimiento exponencial de los artículos publicados sobre el tema, identificación de zonas de Bradford, siendo la revista Medicina Clínica la más productiva. En el año 1995, la investigación *10 años de literatura en SIDA (1983-1992): Análisis bibliométrico*, identifica nuevamente crecimiento exponencial de las publicaciones, ubicadas en el IME, desde 1983, en España y en otros países. Otro trabajo es *Spanish bibliographic production on AIDS. Bibliometric approach*, aplicado en la base de datos española del ISBN, cuyos resultados ubican a Barcelona y Madrid como las ciudades más destacadas, se observa un predominio del idioma español, con un porcentaje de traducciones del inglés.

Entre otros aportes sobre el tema tenemos: *Un decenio de literatura sobre el SIDA* (1999), según la cual, cerca del 60% de revistas indizadas por MEDLINE publicaron al menos un artículo en los últimos 10 años, demostrando así su crecimiento en las actividades de investigación y de control sobre esta enfermedad; *AIDS in Haiti: a bibliometric analysis* (2000), EE.UU., el Haití y Canadá eran los principales productores, y los aspectos publicados estaban centrados en epidemiología y complicaciones del SIDA. *Agenda científica y extracientífica en VIH/SIDA* (2002), se identificaron autores de Brasil y México con mayor número de presentaciones; y *Análisis bibliométrico de la literatura publicada sobre SIDA en el Sur de África* (2003), cuyos resultados fueron producción principal en África del Sur, Zimbawe y Zambia, teniendo al inglés como lengua dominante. La alta productividad estaba influenciada por las universidades e instituciones relacionadas con EE.UU y UK.

1.5. Objetivos de la investigación

El presente estudio tiene como objetivos conocer el estado de la producción científica y su evolución, respecto al SIDA, en el período comprendido entre 1997 y 2003, en la base de datos SCiELO; y con ello determinar la revista de mayor productividad, el idioma de publicación de los artículos, la identificación de los años donde se realizó mayor estudio y la cobertura geográfica que abarcan los artículos analizados.

2. Modelo matemático de Bradford

Este indicador bibliométrico de dispersión de la información científica, postula que, dentro de una producción científica determinada existe un núcleo de revistas utilizado con más frecuencia por los autores, y varios grupos o zonas (zonas de Bradford) que incluyen el mismo número de artículos que el núcleo. Así, el número de revistas en las diversas zonas de Bradford cumplirán la relación $1: n: n^2$... Este enunciado nos permite, determinar qué revistas son la más utilizadas por los científicos en una disciplina determinada.

2.1. Limitaciones de la ley de Bradford

La limitación existente sobre esta ley es que Bradford nunca manifestó en su obra, dividir la lista de revistas en más de tres grupos, llevando a muchos autores a trabajar con más zonas, y cuanto mayor es la cantidad de zonas el multiplicador es menor y menos constante. Otra crítica es la formulada por E. Wilkinson, quien plantea “la disparidad entre la formulación verbal y la gráfica por parte del propio Bradford, a pesar de que éste pensara equivocadamente que coincidían”²; a ello se suma el cuestionamiento de no considerar la frecuencia de publicación y los años que participan con el período de la investigación, (Urbizagástegui, 1996), ambas variables que influirán en la productividad de una revista, lo que puede ser la causa que un título de revista se coloque en el tope o al final de la distribución de Bradford. Por lo tanto, la ley, en este campo, no considera a las revistas con menos años de publicación y menos frecuencia y sobreestima aquellas revistas con mayor frecuencia de publicación.

Estas consideraciones ponen de manifiesto que la ley de Bradford, es un indicador defectuoso de la productividad por no considerar la frecuencia, ni el número de años con que una revista participa en el período de publicación.

3. Método de investigación

3.1. Fuente

SciELO³ *Scientific Electronic Library Online* (Biblioteca Científica Electrónica en Línea) es un modelo para publicación electrónica en países en desarrollo. Esta iniciativa tiene por objetivo brindar acceso adecuado y actualizado a la información científico-técnica que se difunde en las revistas científicas de los países en desarrollo, las cuales enfrentan graves barreras de distribución y diseminación.

Especialmente desarrollado para responder a las necesidades de la comunicación científica en los países en desarrollo, y particularmente de América Latina y el Caribe, el modelo proporciona una solución eficiente para asegurar la visibilidad y el acceso universal a su literatura científica, contribuyendo a la superación del fenómeno conocido como “ciencia perdida”. Además, el modelo SciELO contiene procedimientos integrados para la medida del uso y del impacto de las revistas científicas.

3.3. Población, muestra

Para el análisis del estudio bibliométrico de los registros de los artículos de revistas de SIDA se tomarán en cuenta los publicados a partir de 1997, año que indica el funcionamiento de la base de

² López Piñero, 1972. El análisis estadístico y sociométrico de la literatura científica

³ Modelo SciELO. <http://www.scielo.org>

datos SciELO. Por lo que dicho estudio sólo seleccionará los datos que se encontraron en un primer momento de la búsqueda para obtener los datos correspondientes y realizar la ejecución del análisis cuantitativo y posterior interpretación.

Los datos observados en la base de datos SciELO son el número total de documentos que representa nuestra población y para efecto del mencionado estudio se ha seleccionado un total de 315 artículos cuyo número representa la muestra, conformada por diversas revistas cuyos contenidos se basan en estudios de SIDA.

3.2. Método empleado

Se realizó la búsqueda en la base de datos SciELO, utilizando como término de indización AIDS. La descarga de documentos comprendió dos pasos: ingresar a la base de datos y obtener de cada página el total de registros, para su posterior análisis bibliométrico. A fin de conseguir el total de títulos de revistas y, de esta forma, efectuar la cuantificación, se utilizó los programas WordPad⁴ y Excel⁵. Asimismo, en la hoja de cálculo Excel se confeccionó la tabla con los datos y funciones correspondientes para cada título de revista, lo que permitió calcular la distribución, en orden de productividad decreciente de artículos de publicaciones periódicas de acuerdo al modelo matemático de Bradford (Gorbea Portal, Salvador, 1996).

La muestra analizada estuvo formada por 315 registros de artículos de revistas de la base de datos SciELO, en el período comprendido entre 1997 y mayo de 2003. Dichos artículos fueron previamente depurados, para la correspondiente tabulación, análisis e interpretación de los datos, ya que algunos estudios no trataban sobre SIDA.

A través de una tabla confeccionada en Excel, se elaboró la distribución decreciente de artículos de revistas, en la que se identificó las zonas en que se dividen las revistas de acuerdo con la cantidad de artículos que aporta cada una. La tabla se contó con las columnas siguientes⁶:

Columna A: número de títulos de revistas.

Columna B: número de artículos por título de revistas.

Columna C: (A*B) cantidad total de artículos por revista.

Columna D: cantidad acumulada de títulos de revistas = n.

Columna E: cantidad acumulada de artículos = R (n).

Columna F: logaritmo de la cantidad acumulada de títulos de revistas = Log(n).

Esta columna es el resultado del cálculo de cada uno de los puntos que conforman la recta de la representación gráfica del modelo. Para establecer matemáticamente las coordenadas que identifican cada punto de la recta calculada, se emplea el método de los mínimos cuadrados en la función logarítmica: $Y=a+b.X$

⁴ WordPad es una marca registrada de Microsoft Corporation

⁵ Excel es una marca registrada de Microsoft Corporation

⁶ Basado en Gorbea Portal, Salvador 1996. El modelo matemático de Bradford.

donde:

$Y = Rc(n)$ cantidad calculada de artículos.

a = intercepto, punto calculado donde se corta la recta con el eje de las Y .

Cuando este valor es negativo indica que se corta por debajo del eje de la X o abscisa.

b = a la pendiente de la recta, es decir, la altura que alcanza la curva en relación con el eje de las X .

a y b son valores constantes para cada punto que se suma y multiplica por cada valor de $\log X$, es decir del $\log (n) =$ al logaritmo de la cantidad acumulada de títulos. Este procedimiento permitió tabular las más importantes publicaciones en la temática de acuerdo al modelo de Bradford.

Asimismo, por medio de la información obtenida de los registros de los artículos de cada revista, se analizó el idioma de publicación, fecha de publicación y su cobertura geográfica. Para el análisis de cada variable que caracterizan a esta dispersión se emplearon técnicas estadísticas y el uso de gráficos para su demostración.

4. Resultados encontrados

A	B	C = (A X B)	D	E	F	G
TITULOS REVISTAS	ARTICULOS REVISTAS	TOTAL DE ARTICULOS	ACUMULADO REVISTAS N	ACUMULADO ARTICULOS $R(n)$	Log DE REV.ACUM $\log (n)$	CANTIDAD ART.CALC. $Rc(n)$
1	37	37	1	37	0.0000000	24.65938595
1	33	33	2	70	0.3010300	84.04822723
1	31	31	3	101	0.4771213	118.7884723
1	28	28	4	129	0.6020600	143.4370685
1	26	26	5	155	0.6989700	162.556005
1	23	23	6	178	0.7781513	178.1773136
1	19	19	7	197	0.8450980	191.3849418
1	17	17	8	214	0.9030900	202.8259098
1	11	11	9	225	0.9542425	212.9175587
1	6	6	10	231	1.0000000	221.9448463
2	8	16	12	247	1.0791812	237.5661549
2	7	14	14	261	1.1461280	250.7737831
2	5	10	16	271	1.2041200	262.2147511
2	3	6	18	277	1.2552725	272.3064
3	4	12	21	289	1.3222193	285.5140282
7	2	14	28	303	1.4471580	310.1626244
12	1	12	40	315	1.6020600	340.7225289

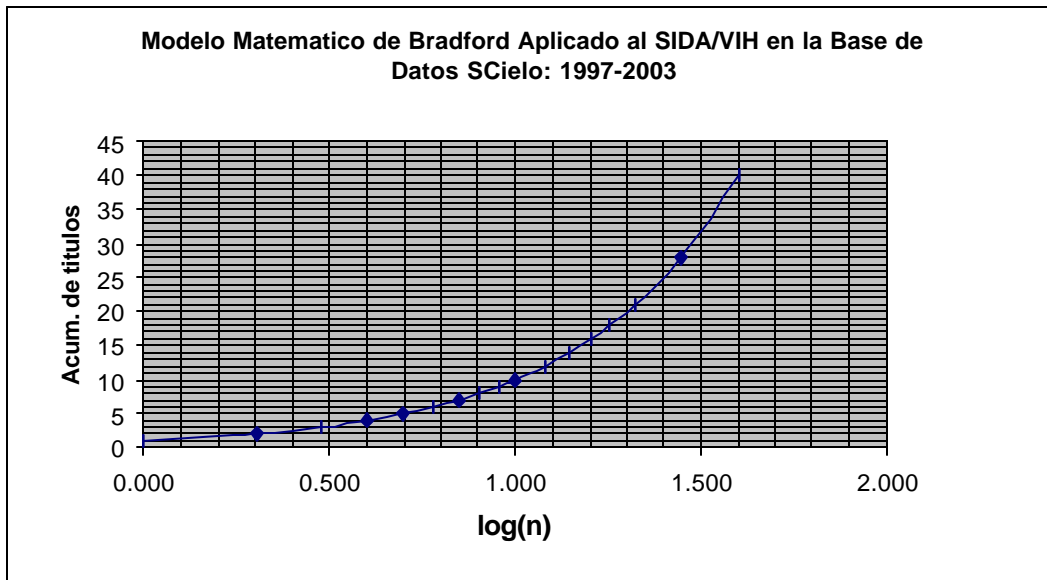
$A =$ 24.65938595

$N =$ 1

2047

$B =$ 197.2854604

$R =$ 0.989400222



En la tabla I observamos los datos analizados de acuerdo a su productividad, se recogieron 315 artículos y 40 títulos de revistas. Para determinar las zonas se dividió la cantidad total de artículos (315), en tres partes (315/3) y las zonas quedaron distribuidas de la siguiente manera:

Tabla I

<i>Núcleo</i>	Revistas	3
	Artículos	101
<i>Zona2</i>	Revistas	7
	Artículos	130
<i>Zona3</i>	Revistas	30
	Artículos	84

La zona nuclear está formada por 3 revistas que contienen, aproximadamente, el 30% del total de información analizada. Se observa la tendencia de concentración de la materia en tres títulos, que son Sociedade Brasileira de Medicina Tropical, Bulletin of the World Health Organization y Cadernos de Saúde Pública, asimismo el incremento de títulos por la llamadas "zonas".

Modelo de Bradford

$$p : p_1 : p_2 : 1 : n : n_2$$

Lo que significa que:

Los títulos de la tercera zona crecen aproximadamente en el cuadrado de los títulos de la razón que hay entre la tercera y la segunda y la segunda y la primera zona.

Reemplazando los datos obtenidos en el presente estudio en el modelo de Bradford, tenemos que: El número promedio de artículos por revistas en las diferentes zonas r , r_1 , r_2 , respectivamente, es igual a:

$$\begin{array}{lcl}
 r = \frac{m}{p} & r_1 = \frac{m_1}{p_1} \text{ y} & r_2 = \frac{m_2}{p_2} \\
 r = \frac{101}{3} & r_1 = \frac{130}{7} \text{ y} & r_2 = \frac{84}{30} \\
 r = 33.67 & r_1 = 18.57 \text{ y} & r_2 = 2.8
 \end{array}$$

es decir, se cumple lo propuesto, que

$$p < p_1 < p_2 < \dots$$

y que de la forma inversa:

$$r > r_1 > r_2 > \dots$$

por otra parte: $pr = p_1 r_1 = p_2 r_2 \dots = m$

haciendo la respectiva sustitución tenemos que:

$p \cdot r = m$	$p_1 \cdot r_1 = m_1$	$p_2 \cdot r_2 = m_2$
$3 \cdot 33.67 = 101.01$	$7 \cdot 18.57 = 129.99$	$30 \cdot 4.67 = 140.1$
$101 = m$	$130 = m_1$	$140 = m_2$

Reemplazando:

$$\begin{array}{ccc}
 \frac{p_1}{p} & \frac{p_2}{p_1} & \\
 \frac{7}{3} & \frac{30}{7} & \frac{2.3 + 4.3}{2} \\
 n = 3.3 & &
 \end{array}$$

n = coeficiente de proporcionalidad de títulos entre zonas o multiplicador de Bradford

Pn = número de títulos de revistas por zonas

Cálculo del multiplicador de Bradford para la muestras analizada: Los títulos observados por zonas son: 3, 7, 30. La constante $n = 3.3$ Sustituyendo la constante n por 3.3 tenemos:

$$\begin{array}{lcl}
 n: & n_1: & n_2 = 1: n: n^2 \\
 (3): & (3): & (3) \\
 1: & 3: & 9 \\
 3(1) & 3(3) & 3(9) \\
 3: & 9: & 27 = 3: 7: 84
 \end{array}$$

Para los indicadores los resultados se presentan a continuación:

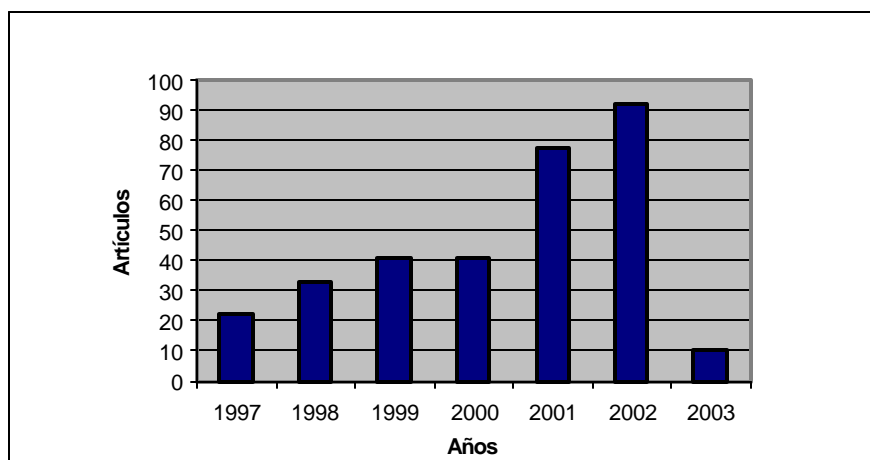


Gráfico N° 1. Distribución por años

El número de documentos sobre el SIDA, es muy incipiente en los primeros años de vida de la base de datos SciELO (1997), y partir de los siguientes años empieza un crecimiento de manera lenta, pero superior a los años anteriores. Este impulso en el crecimiento de los documentos indizados se debe a que la metodología SciELO aún está siendo implantada en esta región geográfica y existen muchas revistas que se están acogiendo a ella. Como podemos observar, en los años 2001 (24%) y 2002 (30%), la producción llega a cubrir la participación total del tema en análisis, como consecuencia de los hechos anteriormente explicados.

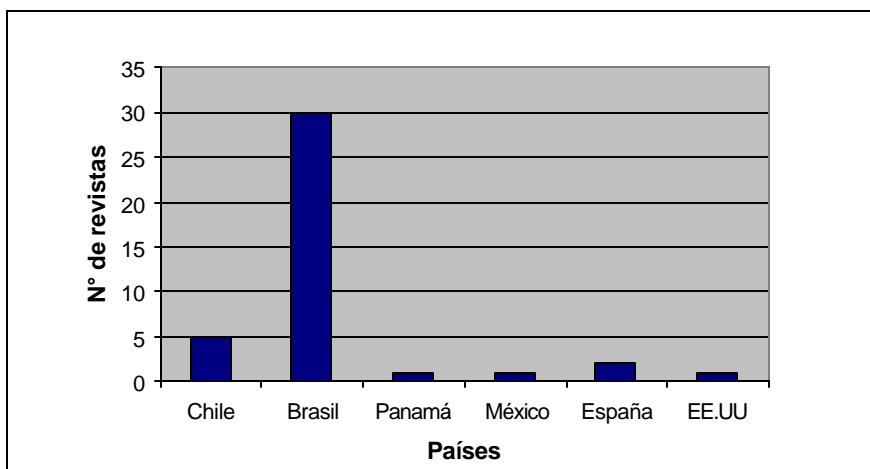
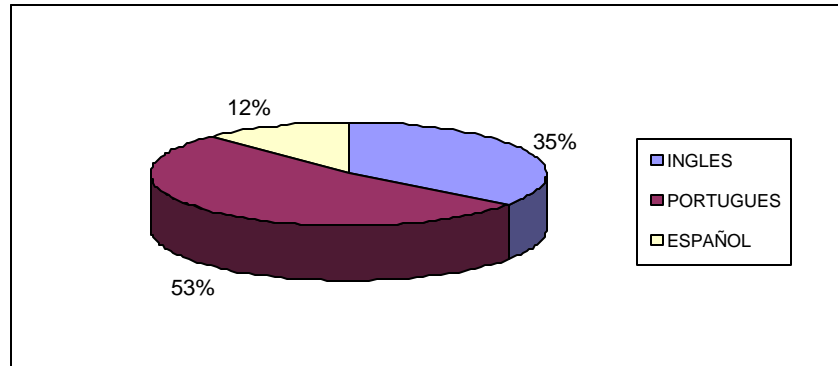


Gráfico N° 2. Distribución por países

Al realizar el análisis de porcentajes totales de la producción por países, para distinguir qué país tiene mayor presencia a nivel latinoamericano y fuera de la región (Europa) de acuerdo a su aporte en la base de datos SciELO.

Como podemos apreciar en el gráfico 2, los resultados no son tan sorprendentes referentes al primer puesto y siguientes: Brasil, Chile. No sorprende porque estos países se encuentran en un nivel muy superior en el área de Ciencias de la Salud, en esta parte de la región. Además, a ellos corresponde la iniciativa de creación de la base de datos en estudio. Éste es un hecho que debe tenerse en cuenta para la participación de los otros países (México, Panamá, España y EE.UU.).

La pequeña sorpresa que nos tiene éste indicador, es la presencia de EE.UU., que muestra una pequeña participación (3%) debido a su reciente acoplamiento a SciELO.



Gráfica N° 3. Distribución por idiomas

De acuerdo con el análisis este indicador permitirá identificar la tendencia de producción en una u otra lengua, ya que se “manifiesta como uno de los factores que dificulta el proceso de comunicación científica y por consiguiente, la posibilidad de que el investigador pueda conocer”⁷. El predominio de los idiomas portugués e inglés, obedece entre otras razones tanto al origen de la base de datos, como a la visibilidad que puedan llegar a tener.

Conclusiones

- En el período comprendido entre 1997 y mayo del 2003 se publicaron 315 artículos de revistas científicas relacionadas al SIDA. Los estudios sobre SIDA publicados en la región de Latinoamérica ha crecido paulatinamente conforme han ido pasando los años.
- Las revistas de mayor productividad que conforman el núcleo son: Revista da Sociedade Brasileira de Medicina Tropical, Bulletin of the World Health Organization y Cadernos de Saúde Pública.
- Los idiomas en que han sido publicados los artículos, ocupando el primer lugar se encuentra el portugués, seguido del inglés y por último el idioma español.
- Los años en que se realizaron más investigaciones fueron el 2002, 2001, 1999 y 1998.
- Los países más productivos son Brasil y Chile, seguidos por México, Panamá, España y EE.UU.

⁷ Gorbea Portal, Salvador. El idioma en la generación y uso de la información ... p. 73.

REFERENCIAS

Cañedo, R.; Hernández, W.; Gutiérrez, A.; Guerrero, L. y Morales, M. (1999). Producción científica de y sobre Cuba procesada en la base de datos Medline en el período 1986-1995. *ACIMED* 7 (2), 104-144.

De la Cueva, A.; Almero A, O. y Gimenez, J. (1992). *10 años de literatura en Sida.(1983-1992) : Análisis bibliométrico.* [En línea] <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=7654834&dopt=Abstract> [Consulta 21 mayo 2003]

Ferreiro Aláez, L. (1984). Dispersiones de la literatura científica: su ajuste a la ley de Bradford. *Revista Española de Documentación Científica*, 7 (2).

Gorbea Portal, S. (1996). El modelo matemático de Bradford: su aplicación a las revistas latinoamericanas de las ciencias bibliotecológica y de la información. México, D. F. : UNAM, Centro Universitario de Investigación Bibliotecológica.

Licea, J.; Valles, J. e Izazola, J (2002). Agenda científica y extracientífica en VIH/SIDA. *Anales de Documentación*, 4, 139-149.

López Piñeiro, J. (1972). El análisis estadístico y sociométricos de la literatura científica.

López López, P. (1996). Bibliometría: la medida de la información. Manual de información y documentación. Madrid : PIRÁMIDE.

López López, P. (1996). Introducción a al bibliometría. Valencia: PROMOLIBRO.

Macias-Chapula, C. AIDS in Haiti: a bibliometric analysis. [En línea] <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=10658964>> [Consulta 21 mayo 2003]

Modelo Scielo. *Metodología SCiELO.* [En línea] <<http://www.scielo.org>> [Consulta 21 mayo 2003]

Oscá, J. [Spanish bibliographic production on AIDS. Bibliometric approach]. [En línea] <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=9424119&dopt=Abstract> [Consulta: 21 mayo 2003]

Piscoya Hermoza, L. (1995). Investigación científica y educacional: un enfoque epistemológico. Lima

Spinak, E. (1996). Diccionario enciclopédico de bibliometría, cienciometría e informetría. Caracas: UNESCO.

Spinak, E. (2001). Indicadores Cienciométricos. *ACIMED*, 9, suplemento 2001.

Urbizágastegui Alvarado, R. (1996). Una revisión crítica de la ley de Bradford. *Investigación Bibliotecológica*, 10 (20), 16-25.

Revistas	Participación
Revista da Sociedade Brasileira de Medicina Tropical	37
Bulletin of the World Health Organization	33
Cadernos de Saúde Pública	31
Revista de Saúde Pública / Journal of Public Health	28
Memórias do Instituto Oswaldo Cruz	26
Revista do Instituto de Medicina Tropical de São Paulo	23
Brazilian Journal of Infectious Diseases	19
Revista Panamericana de Salud Pública / Pan American Journal of Public Health	17
Revista da Associação Médica Brasileira	11
Revista Española de Salud Pública	8
Salud Pública de México	8
Arquivos de Neuro-Psiquiatria	7
Brazilian Journal of Medical and Biological Research	7
Revista Médica de Chile	6
Pesquisa Odontológica Brasileira	5
Revista Brasileira de Psiquiatria	5
Arquivos Brasileiros de Cardiologia	4
Rev. Chil. Infectol.,	4
Sao Paulo Med. J.,	4
Psicol. Reflex. Crit.,	3
Psicol. Usp,	3
An. Acad. Bras. Ciênc.,	2
Arq. Gastroenterol.,	2
Ciênc. Saúde Coletiva,	2
Parasitol. Día,	2
Parasitol. Latinoam.,	2
Rev. Chil. Pediatr.,	2
Rev. Microbiol.,	2
Arq. Bras. Oftalmol.,	1
Estud. Psicol. (Natal),	1
Gac Sanit,	1
Hist. Cienc. Saude,	1
J. Bras. Patol. Med. Lab.,	1
J. Pneumologia,	1
J. Venom. Anim. Toxins,	1
Rev Odontol Univ São Paulo,	1
Rev. Bras. Ginecol. Obstet.,	1
Rev. Bras. Hematol. Hemoter.,	1
Rev. Bras. Otorrinolaringol.,	1
Rev. Latino-Am. Enfermagem,	1

SOBRE LOS AUTORES:

Rudy Richard Chiroque Solano

Lima, Perú

Estudiante de Bibliotecología y Ciencia de la Información. Universidad Nacional Mayor de San Marcos

c.e: richardchiroque@lycos.es

Ariela Jesús Medina Valdivia

Lima, Perú

Estudiante de Bibliotecología y Ciencia de la Información. Universidad Nacional Mayor de San Marcos

c.e: amedinavaldivia@yahoo.es