Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Library & Information Science Research 29 (2007) 124–134
Library & Information Science Research

HIV/AIDS research in India: A bibliometric study

Swapan Kumar Patra [a], Prakash Chand [b,*]

[a] *Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, New Delhi-110016, India*

[b] *National Institute of Science Communication and Information Resources, 14-Satsang Vihar Marg, New Delhi-110067, India*

Available online 9 November 2006

Abstract

This article looks at the growth over time of Indian AIDS research output based on bibliographic data from *PubMed* and *Web of Science*. Authorship distribution was examined using Lotka's law. Bradford's law of scattering was used to identify core journals. The study identifies active institutions and statewide distributions of Indian AIDS research output. The yearly analysis of data shows that there is a rapid growth of literature from 1992 onwards. Still, in an international sense, relative productivity of India is low and requires more focused research and development.

1. Introduction

India, with a population of over one billion, is experiencing a rapid and extensive spread of HIV. According to a UNAIDS estimate, about 5.134 million people are infected with HIV in India (Cohen, 2004). UNAIDS India (2006) has also reported that India has a HIV/AIDS infection rate estimated at 0.91% of the adult population. India initiated HIV-prevention activities in the very early stages of the epidemic and committed to prevention efforts. The National AIDS Control Organization (NACO; 2006) was constituted in 1986 and the National AIDS Control Program launched a year later.

* Corresponding author. Fax: +91 11 26862228.
E-mail addresses: skpatra@gmail.com (S.K. Patra), pchand_insdoc@yahoo.com, prakashc@niscair.res.in (P. Chand).

0740-8188/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.lisr.2006.08.010

NACO is the apex body for policy formulation and implementation to prevent and control HIV/AIDS in India. NACO works in two phases. The first phase, starting from its conception to 1999, mainly stresses public awareness and other preventive measures. In the second phase, research and development (R&D) are the key priorities. To enhance R&D in this area, NACO identifies and promotes Indian academic institutions that conduct epidemiological and operational research for the treatment of HIV/AIDS, including research on indigenous systems of medicine such as ayurveda and siddha. Emphasis is also given to developing an HIV vaccine and promoting strong collaboration among national R&D institutions as well as with international bodies such as the World Health Organization (WHO).

2. Problem statement

To understand India's strength and capability in AIDS research, a bibliometric study was conducted to portray India's research and development profile. Bibliometrics, a term introduced by Pritchard (1969), uses mathematical formulas to quantify productivity distribution and yield of communication output and allows for predicting and studying scientific progress. This study explores the growth of AIDS research in India over time and identifies active researchers and Indian institutions, their productivity, authorship patterns, and core journals.

3. Literature survey

With the high incidence of AIDS cases, growth of R&D literature has been rapid. Systematic monitoring and evaluation of research on HIV/AIDS is becoming essential. This unprecedented growth of literature was the impetus for a scientometric study intended to help scientometrists, information scientists, policy and decision makers, and related research workers. Sengupta and Kumari (1991) analyzed AIDS literature published during 1976–1986 and found that literature on HIV/AIDS was growing exponentially. Self et al. (1989) assessed the growth of AIDS literature from 1982 to 1987 based on MEDLINE, Magazine Index, and National Newspaper Index. Forney (1990) identified a list of 18 core journals of AIDS research. Pratt (1992) used the MRAI database to carry out a bibliometric analysis of AIDS literature from 1981 to 1990. Gullion and Huber (2003) examined literature specific to the use of complementary and alternative medicine where HIV/AIDS is concerned to determine publication patterns. In developing countries, national level descriptive and evaluative bibliometric approaches have been used to study AIDS literature, for example, in Latin America and the Caribbean between 1980 and 1996 (Macías-Chapula et al., 1998), in Haiti (Macías-Chapula, 2000), in Southern Africa (Macías-Chapula et al., 2002), in Sub-Saharan Africa (Macías-Chapula & Mijangos-Nolasco, 2002), and in Uganda and Kenya (Ocholla & Onyancha, 2004). A bibliometric study on AIDS/HIV research in India is not available.

4. Methodology

The data for this study were collected from the National Library of Medicine's (NLM) *PubMed* and the Institute for Scientific Information's (ISI) *Web of Science* for 1988–2005. The *PubMed* online bibliographic database, available free of cost, *PubMed* scans journals in 40 different languages related to biomedical science. *Web of Science* is a multidisciplinary bibliographic database that provides information from approximately 8,700 international journals and is used to map worldwide science and technology (S&T) data. Data for the period 1988–2005 was downloaded using EndNote7 software, developed by Thomson ISI Research Soft. The search terms used to retrieve data were "[Acquired Immunodeficiency Syndrome OR AIDS OR HIV and Publication place/Author's affiliation=(India)]." Acquired Immunodeficiency Syndrome and HIV are MeSH subject heading terms (Brooks et al., 1990) and were used to retrieve maximum and relevant records from the databases. The total number of collected records was 2,718 (*PubMed* 1,173; *Web of Science* 1,545). The data consist of bibliographic information (i.e., author, title of the article, name of journal, place of publication, type of document, year of publication, and keywords). Since the data were from two sources, 540 records appeared as duplicates and were removed using EndNote's duplicate removal application. After removing duplicate records, 2,178 unique records remained.

In the case of multinational collaborations, only the Indian author's address was taken as the author's address. Additionally, the first author's address was taken for affiliation because *PubMed* only indexes the first author's address. *Journal Citation Reports* (*JCR*) (2004) was used to determine impact factors of journals.

5. Results

At first glance, analysis of the data shows that most Indian researchers publish in the English language (e.g., all found articles from *PubMed* and 1,543 articles from *Web of Science* are written in the English language). In addition, Indian researchers mainly depend on communications appearing in journals: *Web of Science* results show 1,180 journal articles, 46 conference proceedings, 129 letters, 32 editorials, 137 reviews, and 21 others; *PubMed* results show 1,168 journal articles.

5.1. Literature growth

This study's results contain 2,178 records covering journal articles, case reports, and review articles on Indian AIDS research. Fig. 1 shows the growth of Indian AIDS literature. Initially, growth was slow but gradually picked up over time. From 1992 onwards, the growth of literature is exponential, and in last two years over 300 articles a year have been published. Fig. 2 shows that the highest number of publications is from the United States followed by the United Kingdom and France. China is ranked below India.

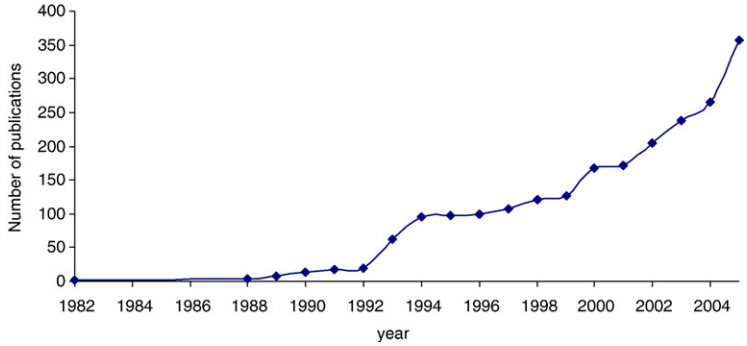

Fig. 1. Growth pattern.

5.2. Authorship Pattern

During 1982–2005, 5,071 authors produced 2,178 publications of Indian origin, with an average of 2.3 authors per article. The most productive author is S. Solomon, from the Madras Medical College and Government General Hospital, Chennai with 68 publications, followed by N. Kumarasamy from the Christian Medical College (CMC), and Vellore, Tamil Nadu, with 55 publications. A list of authors who produced more than 20 publications is given in Table 1. The author data indicate that the R&D activity on HIV/AIDS gained momentum from 1990s.

Lotka's law describes the frequency of publication by the authors in a given subject (Lotka, 1926; Nicholas, 1980; Pao, 1985). To validate Lotka's law, a log-log plot was drawn between the percentage of authors and the percentage of their publications (Fig. 3). Using Lotka's law, the value of constant C and n are determined and found to be 0.66 and −2.19, respectively. A Kolgomorov–Smirnov test (Newby et al., 2003; Pao, 1985) shows that literature is not following the original Lotka distribution. In Indian HIV/AIDS research output, about 3,419 (67.4%) authors have one publication, 890 (17.5%) have two, and 286 (5.6%) authors have three. Authors with more than 10 publications account for only 1.5% (91 authors).

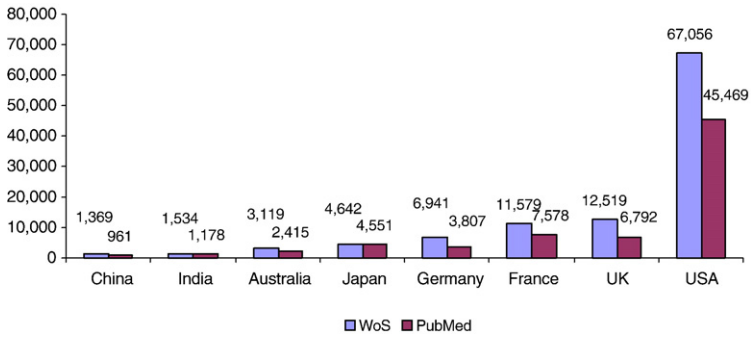

Fig. 2. Distribution of literature by country (the coverage includes up to December 2005; duplicates are not removed).

Table 1
Productive authors

Name of author	Number of publications	Productive years
1. Solomon, S.	68	1994–2005
2. Kumarasamy, N.	55	1995–2005
3. Mehendale, S.	48	1993–2005
4. Singh, S.	44	1989–2004
5. John, T. J.	36	1994–2005
6. Kumar, A.	35	1992–2005
7. Bollinger, R. C.	31	1995–2005
8. Paranjape, R. S.	28	1998–2005
9. Bhattacharya, S. K.	28	1991–2005
10. Seth, P.	27	1999–2005
11. Sridharan, G.	27	1999–2005
12. Kannangai, R.	26	1990–2002
13. Babu, P. G.	26	2002–2005
14. Mayer, K. H.	25	1995–2004
15. Biswas, J.	25	1999–2005
16. Ramalingam, S.	25	1994–2005
17. Chattopadhya, D.	24	1992–2005
18. Panda, S.	23	1992–2005
19. Madan, A. K.	22	2001–2005
20. Kumar, S.	21	2000–2005
21. Sharma, S. K.	21	1995–2005

5.3. Identification of core journals

Altogether there are 784 journals represented by the 2,178 articles. Out of these, 461 journals (58 %) published one article, 149 journals (19%) published two articles, and 50

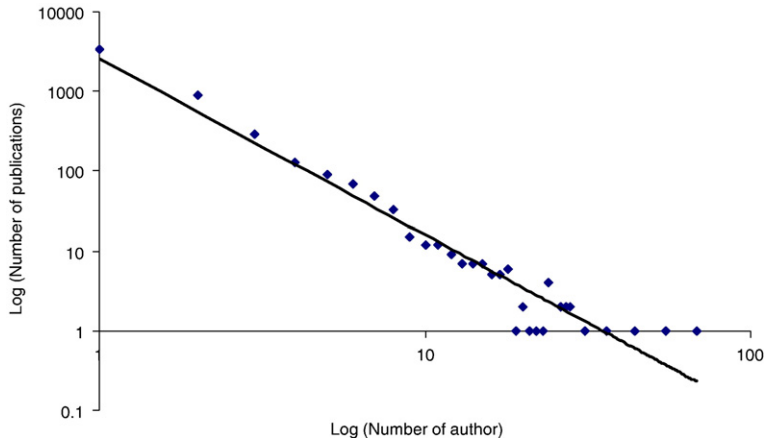

Fig. 3. Authorship pattern.

Table 2
Journals with number of publications, impact factor, and publisher

Name of journal	Number of publications	Impact factor (2004)	Place, publisher
1. Indian Journal of Medical Research	118	0.600	New Delhi, Indian Council of Medical Research
2. National Medical Journal of India	60	0.626	New Delhi, All India Institute of Medical Sciences
3. AIDS	48	5.893	USA, Lippincott Williams and Wilkins
4. International Journal of STD and AIDS	33	1.506	UK, Royal Society of Medicine
5. Journal of Acquired Immune Deficiency Syndromes	31	–	USA, Lippincott Williams and Wilkins
6. Lancet	30	21.713	Amsterdam, Elsevier
7. Indian Journal of Pediatrics	29	–	New Delhi, All India Institute of Medical Sciences
8. AIDS Research and Human Retroviruses	28	2.375	New York, Mary Ann Liebert
9. Indian Pediatrics	27	–	New Delhi, Indian Academy of Pediatrics
10. The Journal of communicable Diseases	26	–	Delhi, Indian Society for Malaria and other Communicable Diseases
11. Journal of the Association of Physicians of India	23	–	Bombay, Association of Physicians of India
12. Tetrahedron Letters	23	2.484	Amsterdam, Elsevier
13. Bioorganic and Medicinal Chemistry	21	2.018	Amsterdam, Elsevier
14. Journal of Clinical Microbiology	20	3.439	Washington, American Society for Microbiology
15. Neurology India	18	0.339	Bombay, Neurological Society of India
16. Indian Journal of Chest Disease and Allied Sciences	17	–	New Delhi, Vallabhbhai Patel Chest Institute
17. International Journal of Dermatology	16	0.884	UK, Blackwell Publishing
18. Indian Journal of Ophthalmology	16	–	Bombay, All India Ophthalmological Society
19. Sexually Transmitted Infections	16	2.204	London, BMJ Publishing Group
20. Asian Journal of Chemistry	14	0.262	Gaziabad, R.A. Agarwal
21. Current Science	13	0.688	Bangalore, Indian Academy of Sciences
22. Postgraduate Medical Journal	13	0.807	London, BMJ Publishing Group
23. Acta Cytologica	13	0.831	USA, Science Printers and Publishers, St. Louis
24. Tetrahedron	13	2.643	Amsterdam, Elsevier
25. Indian Journal of Chemistry Section B	13	0.476	New Delhi, Council of Scientific and Industrial Research
26. Clinical Infectious Diseases	12	5.594	USA, University of Chicago Press
27. Journal of Infectious Diseases	12	4.943	USA, University of Chicago Press
28. AIDS Care	11	1.600	UK, Taylor and Francis
29. International Journal of Tuberculosis and Lung Disease	11	1.484	Paris, International Union Against Tuberculosis and Lung Disease

journals (6 %) published three articles. Twenty-nine journals (4%) published about one third of the total (see Table 2). The remaining two-thirds of the articles are scattered among 401 journals. Bradford's (1934) law of scattering is employed here to identify "core" journals. In a Bradford plot, the cumulative total of publications is plotted against the logarithm of a journal's rank (Fig. 4). Core journals are those that lay on the initial curved part of the "S"-shaped Bradford plot until it tangentially becomes a straight line. Here, in Fig. 4, we observe that the slope of the curve decreases slightly after the 29th journal, so it appears that about 29 journals form a core.

5.4. Institutions

The list of institutions is quite long as there are 298 institutes that have published 2,178 articles. The All India Institute of Medical Sciences (AIIMS) in Delhi is the most productive, with 204 (9.3%) publications, followed by the Christian Medical College (CMC), Vellore, with 106 (4.2%) publications and the National AIDS Research Institute, Tamilnadu, with 7 (3.2%). The Indian Council of Medical Research (ICMR) sponsored projects in different institutions and ranks fourth with 61 (2.8%) publications. The top 10 institutions have published about 30% of total literature (Table 3). The remaining 70% of the articles are scattered among 264 institutions.

5.5. States

The addresses of the first authors were taken to study author affiliation, and in the case of multinational collaborations, only Indian authors' addresses are considered. Anonymous addresses were discarded. Delhi is the most productive among the metropolitan cities. Delhi has a total of 528 (24.2%) publications from the 47 institutions followed by Mumbai with 247 (12%) publications from 35 institutions, Chennai with 183 (8%) from 25 institutions, and Kolkatta with 127 (6%) publications from 15. Delhi is also at the top among the states (Table 4).

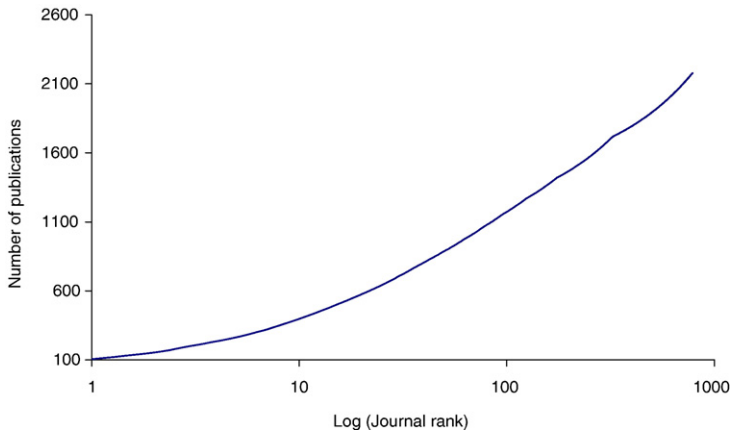

Fig. 4. Bradford plot shows the journal ranks and their cumulative number of publications.

Table 3
Productive institutions

Name of institute	Number of publications	City/state	Percentage of total
1. All India Institute of Medical Sciences	204	Delhi	9.3
2. Christian Medical College	106	Vellore	4.8
3. National AIDS Research Institute	70	Pune	3.2
4. ICMR projects	61	Different locations	2.8
5. Indian Institute of Technologies	61	Different locations	2.8
6. Post Graduate Institute of Medical Education and Research	59	Chandigarh	2.7
7. Y.R.G. Centre for AIDS Research and Education	44	Chennai	2.0
8. Indian Institute of Chemical Technology	40	Hyderabad	1.8
9. University of Delhi	40	Delhi	1.8
10. Banaras Hindu University	36	Varanasi	1.6

Among Union territories, Chandigarh and Pondicherry have high institution/publications ratios. Three institutions from Chandigarh contributed 113 papers, and the Jawaharlal Institute of Post Graduate Medical Education and Research in Pondicherry contributed 26 papers.

Table 4
State distribution of AIDS literature

States	Number of articles	Number of institutions	Publications/ institutions	Percentage of total
Delhi	528	47	11.2	24.2
Maharastra	408	60	6.8	18.7
Tamil Nadu	300	42	7.1	13.7
Uttar Pradesh	135	13	10.3	6.1
Karnataka	125	26	4.8	5.7
Andhra Pradesh	104	17	6.1	4.7
Chandigrah	113	3	37.6	5.1
West Bengal	118	17	6.9	5.4
Kerala	51	15	3.4	2.3
Rajasthan	34	5	6.8	1.5
Gujarat	23	9	2.5	1.0
Pondicherry	26	1	26.0	1.1
Manipur	28	9	3.1	1.2
Punjab	38	3	12.6	1.7
Madhya Pradesh	7	8	0.8	0.3
Goa	8	3	2.6	0.3
Bihar	11	2	5.5	0.5
Haryana	28	2	14.0	1.2
Orissa	10	2	5.0	0.4
Jammu and Kashmir	3	2	1.5	0.1
Himachal Pradesh	3	1	3.0	0.1
Assam	3	1	3.0	0.1
Others	74	10	–	3.3

6. Discussion

The growth of HIV/AIDS research output indicates the stabilization of R&D efforts in the country from the early 1990s. Although there is an identifiable growth in R&D output for the 24-year span of time studied here, there are still problems with the availability of well-trained people, the number of institutions involved in R&D, and the number of nationwide AIDS control programs.

The data show that there are only 21 authors who have more than 20 publications to their credit. Relative to the size of the country, population, availability of S&T manpower, countrywide AIDS research campaigns, and research programs, a figure of 21 productive authors is insignificant and reflects a lack of widespread involvement of researchers.

The publication pattern in journals shows that there is a tremendous scattering of AIDS literature. Scattering of literature poses a great problem in information retrieval in general and affects precision value while retrieval output is taken for a given problem (Ming et al., 2000; Sittig, 1996; Ungern-Sternberg, 2000).

Bradford analysis can reveal core journals, which represent one-third of the total literature in a given field. Based on core journal analysis, it is possible to identify journals that represent specific percentages of articles of a given subject, and this is useful in collection development. The application of Bradford analysis has gained more relevance and importance due to the increasing costs of R&D journals. Among the core journals identified in this study, 12 are published in India and 6 are not covered in the *JCR*. It is also important to note that the curve took a "J" Shape rather than a typical "S" shaped sigmoid curve. There is no "gross droop" at the end. After an initial rise, the graph has taken almost a linear shape, which means the subject field is narrow and growing (Brooks, 1968, 1969; Ungern-Sternberg, 2000).

Productivity-wise, public-funded educational institutions are dominant in comparison to public-funded R&D laboratories, though infrastructure and research facilities are better in the latter. The reason for better performance in the higher educational institutes, when compared to national laboratories, is the presence of a large number of doctoral students (Anurachalam, 1999; Patra & Chand, 2005). Delhi is probably the most productive metropolitan city because of high-profile organizations such as the All India Institute of Medical Sciences, the National Institute of Immunology, Delhi University, and also being the capital of India.

7. Conclusion

HIV is one of the most serious diseases affecting India and may have a major impact on the overall spread of HIV in Asia and the Pacific. The research output and its growth indicate that research activities in this field are still in their infancy. With its large population, infrastructure, and facilities, Indian research output in terms of scientific papers needs more focused R&D efforts to mark a comparable global presence. National-level AIDS programs need an R&D orientation to improve research output. The growing threat of HIV demands a shift from theoretical to R&D-intensive solutions. There is a need to spearhead an HIV/AIDS vaccine development program in India. The spread of India's research programs on HIV/AIDS needs

the involvement of a larger population horizontally covering all the states. Nationwide involvement of academic and R&D institutions will yield better research output and R&D solutions.

This is a preliminary study on HIV/AIDS research output in India; it may trigger more bibliometric studies or indicators for the purpose of evaluating HIV/AIDS research in the country. Such studies would be helpful in carving out appropriate policies to curb the spread of HIV/AIDS in India. Data analysis on HIV/AIDS from all key databases would also contribute to the evolution of a better research strategy for the country.

References

Anurachalam, S. (1999). Mapping life science research in India: A profile based on BIOSIS 1992–1994. *Current Science*, *76*, 1191−1203.

Bradford, S. C. (1934). Sources of information on specific subjects. *Engineering*, *3550*, 85−86.

Brooks, B. C. (1968). The derivation and application of the Bradford–Zipf distribution. *Journal of Documentation*, *2*, 247−265.

Brooks, B. C. (1969). Bradford's law and the bibliography of science. *Nature*, *224*, 953−956.

Brooks, T. A., Bierbaum, E. G., & Brooks, R. M. (1990). Journal frequency analysis of the literature of acquired immune deficiency syndrome (AIDS), 1987–1989. *Information Processing and Management*, *28*, 98.

Cohen, J. (2004). HIV/AIDS: India's many epidemics. *Science*, *304*, 504−509 (April 23).

Forney, C. D. (1990). The acquired immune deficiency syndrome: A bibliometric analysis: 1980–1984. *Science and Technology Libraries*, *10*(4), 45−90.

Gullion, J. S., & Huber, J. T. (2003). Complementary and alternative medicine as represented in the HIV/AIDS body of knowledge: A bibliometric analysis. *Medical Reference Services Quarterly*, *22*(3), 23−32.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Science*, *16*, 317−323.

Macías-Chapula, C. A. (2000). AIDS in Haiti: A bibliometric analysis. *Bulletin of the Medical Library Association*, *88*(1), 56−61.

Macías-Chapula, C. A., & Mijangos-Nolasco, A. (2002). Bibliometric analysis of AIDS literature in central Africa. *Scientometrics*, *54*(2), 309−317.

Macías-Chapula, C. A., Mijangos-Nolasco, A., EguíaLis-Márquez, J. C., Ramírez Godoy, M. E., & García-Hernández, G. (2002). Análisis bibliométrico de la literatura publicada sobre SIDA en el Sur de África. Bibliometric analysis on AIDS publications in South Africa. *Revista Espanola de Documentacion Cientifica*, *25*(4), 433−441.

Macías-Chapula, C. A., Rodea-Castro, I. P., & Narvaez-Berthelemot, N. (1998). Bibliometric analysis of AIDS literature in Latin America and the Caribbean. *Scientometrics*, *41*(1–2), 41−49.

Ming, Y., Shiow-Jem, J., & Sheau-Shin, M. (2000). Bibliomertic study of semiconductor literature, 1978–1997. *Scientometrics*, *49*, 491−509.

National AIDS Control Organization (NACO). 2006. *National AIDS Control Organisation*. Retrieved December 2, 2005, from http://www.nacoonline.org

Newby, G. B., Greenbarg, J., & Jones, P. (2003). Open source software development and Lotka's law: Bibliometric pattern in programming. *Journal of the American Society for Information Science and Technology*, *54*, 169−172.

Nicholas, P. T. (1980). Emperical velidation of Lotka's law. *Information Processing and Management*, *22*, 417−419.

Ocholla, D. N., & Onyancha, O. B. (2004). A comparative study of the literature on HIV/AIDS in Kenya and Uganda: A bibliometric study. *Library and Information Science Research*, *26*, 434−447.

Pao, M. L. (1985). Lotka's law: A testing procedure. *Information Processing and Management*, *21*(4), 305−320.
Patra, S. K., & Chand, P. (2005). Biotechnology research profile of India. *Scientometrics*, *63*(3), 563−593.
Pratt, G. F. (1992). A decade of AIDS literature. *Bulletin of the Medical Library Association*, *80*, 380−381.
Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation*, *25*, 348−349.
Self, P. C., Filardo, T. W., & Lancaster, F. W. (1989). Acquired immunodeficiency syndrome (AIDS) and the epidemic growth of its literature. *Scientometrics*, *17*(1–2), 49−60.
Sengupta, I. N., & Kumari, L. (1991). Bibliometric analysis of AIDS literature. *Scientometrics*, *20*(1), 297−315.
Sittig, D. F. (1996). Identifying a core set of medical informatics serials: An analysis using the MEDLINE database. *Bulletin of the Medical Library Association*, *84*, 200−204.
UNAIDS India. (2006). *HIV epidemic in India*. Retrieved February 14, 2006, from http:// www.unaids.org
Ungern-Sternberg, S. V. (2000). Bradford's Law in the context of information provision. *Scientometrics*, *49*, 161−186.