

CLINICAL SCIENCE

ASSESSING THE SCIENTIFIC RESEARCH PRODUCTIVITY OF A BRAZILIAN HEALTHCARE INSTITUTION: A CASE STUDY AT THE HEART INSTITUTE OF SÃO PAULO, BRAZIL

Beatriz Helena Tess,^I Sérgio Shiguemi Furuie,^{II} Regina Célia Figueiredo Castro,^{III} Maria do Carmo Cavarette Barreto,^{IV} Moacyr Roberto Cuce Nobre^V

doi: 10.1590/S1807-59322009000600013

Tess BH, Furuie SS, Castro RCF, Barreto MCC, Nobre MRC. Assessing the scientific research productivity of a Brazilian healthcare institution: a case study at the Heart Institute of São Paulo, Brazil. Clinics. 2009;64(6):571-6.

INTRODUCTION: The present study was motivated by the need to systematically assess the research productivity of the Heart Institute (InCor), Medical School of the University of São Paulo, Brazil.

OBJECTIVE: To explore methodology for the assessment of institutional scientific research productivity.

MATERIALS AND METHODS: Bibliometric indicators based on searches for author affiliation of original scientific articles or reviews published in journals indexed in the databases Web of Science, MEDLINE, EMBASE, LILACS and SciELO from January 2000 to December 2003 were used in this study. The retrieved records were analyzed according to the index parameters of the journals and modes of access. The number of citations was used to calculate the institutional impact factor.

RESULTS: Out of 1253 records retrieved from the five databases, 604 original articles and reviews were analyzed; of these, 246 (41%) articles were published in national journals and 221 (90%) of those were in journals with free online access through SciELO or their own websites. Of the 358 articles published in international journals, 333 (93%) had controlled online access and 223 (67%) were available through the Capes Portal of Journals. The average impact of each article for InCor was 2.224 in the period studied.

CONCLUSION: A simple and practical methodology to evaluate the scientific production of health research institutions includes searches in the LILACS database for national journals and in MEDLINE and the Web of Science for international journals. The institutional impact factor of articles indexed in the Web of Science may serve as a measure by which to assess and review the scientific productivity of a research institution.

KEYWORDS: Bibliometrics; Medical Research; Scientific production indicators; Cardiology; Brazil.

^I Departamento de Medicina Preventiva, Faculdade de Medicina da Universidade de São Paulo - São Paulo/SP, Brasil.

^{II} Serviço de Informática, Instituto do Coração (InCor), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo - São Paulo/SP, Brasil.

^{III} Centro Latino-Americano de Informação em Ciências da Saúde (BIREME/OPAS/OMS) - São Paulo/SP, Brasil.

^{IV} Serviço de Biblioteca, Documentação Científica e Didática, Instituto do Coração (InCor), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo - São Paulo/SP, Brasil.

^V Unidade de Epidemiologia Clínica, Instituto do Coração (InCor), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo - São Paulo/SP, Brasil.

Email: beatriz.tess@usp.br

Tel: 55 11 3061.7082

Received for publication on February 26, 2009

Accepted for publication on March 09, 2009

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Healthcare, teaching and research are major components of the scientific production process in the healthcare field. Health service users are often invited to take part in research. On the other hand, graduate students and their supervisors comprise the majority of the clinical staff and researchers at health-related university institutions. The interconnected relationship between healthcare, teaching and research activities in the academic environment makes the task of assessing the institutional performance of each of these three components complex.

Monitoring the results of these activities is essential for

formulating, reviewing and improving institutional research policies aimed to assure the appropriate use of financial, human and material resources and to promote strengthening and growth of an institution.¹⁻³ Furthermore, it is necessary to assess the results of scientific studies conducted with public funds in order to assure that the results benefit society.

Bibliometric indicators have been widely used as parameters to evaluate the individual scientific production of researchers or to assess publications in specific thematic fields, as the institutions' scientific contributions are not often analyzed as a whole. Several studies have investigated universities' scientific production,³⁻⁶ but there are few published studies aimed at assessing the institutional scientific production in teaching hospitals.⁷⁻⁹ The Impact Factor (IF) created by Thomson Scientific (formerly ISI – Institute for Scientific Information) has been the bibliometric indicator most frequently used for assessing research output.^{1,10,11}

The Heart Institute (InCor) is one of the 13 institutes of the University of São Paulo Medical School's hospital (HC-FMUSP). It was founded in 1978 and, since its establishment, the Institute performs care, teaching and research activities daily and has become a reference center for cardiology and cardiac surgery. Scientific research in physiotherapy, psychology, information technology, and bioengineering, among other fields of knowledge, is performed at the Institute.

The need to systematically review and assess the performance of the research areas at InCor prompted this study, of which the main objective was to explore a methodology for the assessment of the scientific production of research institutions based on search strategies and an analysis of scientific articles published in journals that are indexed in bibliographic databases.

MATERIALS AND METHODS

In this study, assessment of InCor scientific production comprised retrieval, identification and classification of articles published from January 2000 to December 2003. This time period was chosen because one of the study databases, the Latin American and Caribbean Literature on Health Sciences (LILACS), began to include authors' affiliations in 2000. In addition, a four-year period was deemed to be sufficient to provide data for the analyses necessary to satisfy the objectives of the study.

The retrieved records were analyzed according to the scientific journals in which they were published, databases in which they were indexed, modes of access and number of citations. The InCor publications were retrieved from the international databases Web of Science,

MEDLINE and EMBASE, as well as from the regional databases LILACS and Scientific Electronic Library Online (SciELO). These databases were selected because they were the most comprehensive in the field of life sciences, primarily in regards to health. The database of InCor scientific production, which is maintained by the Scientific Documentation Service of the institution, was used as a control. All articles retrieved from the electronic databases were manually compared to the articles in the institutional database, which enabled identification of flaws in the retrieval strategies and completion of the InCor database.

We considered national journals to be all those published in Brazil, independent of their circulation (national or international), and considered international journals to be those published in other countries. The database indexing indicators and mode of access were analyzed separately based on these categories.

All original articles and reviews retrieved from at least one of the databases mentioned above were included in this study. Proceedings of congresses, editorials, letters to the editor, case reports, brief communications and comments were excluded from the analysis because the content of these sections varies among journals and may include education-oriented publications, such as clinical and radiological discussions. The search strategy was applied to the author affiliation field using the various names of the institution - InCor, *Instituto do Coração da Universidade de São Paulo* and Heart Institute of São Paulo. Articles containing at least one author from InCor were retrieved.

Initially, the results retrieved from each database were considered to represent the institutional production. Articles indexed in more than one database were counted only once. Later the results were compared to the institutional database information. The articles published in indexed journals that were not retrieved through the initial strategy were then manually checked and added to the worksheet after confirmation that at least one author was affiliated with InCor. This strategy was necessary primarily because it is possible to retrieve affiliation only of the first author in the MEDLINE database.

Regarding modes of access, the journals were classified as: i) no electronic access to the full text, including journals that have their own site but provide no access to articles; ii) controlled access, including journals available through the Capes Portal of Journals and; iii) open access, including the SciELO journals.

Categorizing the articles according to their impact took into account the indicators of the Web of Science and the Journal Citation Reports (JCR) in 2005, that is, the number of citations of the articles and the impact factor (IF) of the journals. The IF for a journal in year N is the ratio between

the number of citations of the articles published in N-1 and N-2 and the number of articles published in the same two years. The IF expresses, on average, the frequency that an article published in the journal in year N is cited in the following two years.

Keeping in line with the definition of the IF, we defined the impact factor of an article (article IF) as the mean number of citations of an article in the two years following its publication.

The following formula represents this factor: article IF = $(c_1 + c_2)/2$, where c_1 and c_2 represent the number of citations in the first and second years after publication, respectively. For instance, if an article was published in the year 2000 in the journal indexed in the JCR with two and four citations in 2001 and 2002, respectively, the article IF will be three. This indicator enables comparison of the IF of one article to the IF of journals.

To estimate the impact of an institution on the universe of publications of a given year, we defined the institutional impact factor (institutional IF) as the sum of the article IF of all publications of the JCR of the institution in the year. This indicator allows for comparison with other institutions and groups and takes the size of the corporation into account in the determination of its impact.

RESULTS

The first search based on affiliation of authors in the five databases included in the study retrieved 1253 records, of which 653 were excluded because they were publication types out of the scope defined for this study. Through a manual review of the data available at the Scientific Documentation Service of InCor and comparison with the results obtained through an automated search, four more articles were identified and included in the study; they had been published in journals not indexed in the bibliographic databases. Approximately one third of the 600 articles and reviews retrieved by automated searches were not registered in the institutional database.

The study comprised 604 single records of articles

in journals with at least one author affiliated with InCor during the period between 2000 and 2003. Of this total, 246 (41%) were published in national journals and 358 (59%) in international journals (Table 1). Out of the international journal articles, 325 (91%) were indexed in the Web of Science, MEDLINE and EMBASE databases. This finding was evident in the complementary analysis, since the articles by InCor authors were not always retrieved from the MEDLINE database, which records only the affiliation of the first author. Of the articles published in national indexed journals, 115 (47%) were present in four or more databases, 25 (10%) in three, 45 (18%) in two, and 60 (24%) only in the LILACS database. Four articles were not indexed in any of the selected databases – three of these were published in international journals and one in a national journal.

The scientific production of InCor is better represented in MEDLINE, where 81% of the retrieved articles were indexed. For the productions published in national journals, the LILACS database was more efficient; out of 246 articles, 196 (80%) were found. For international journals, the search was more efficient in the Web of Science; out of 358 articles, 322 (90%) were retrieved with no need for a manual search.

Of the 604 articles, journals with electronic access predominated: 91% (223/246) of articles published in national journals had online access, as well as 97% (346/358) of articles published in international journals (Table 2). Of the national journals, 90% (221/246) were open-access, either through SciELO or through their own sites. The opposite situation occurred in international journals; 93% (333/358) had controlled access. Regarding international titles, 62% (223/358) were identified in the Capes Portal of Journals, whereas 74% (181/246) of the articles published in national journals were in SciELO.

Considering only the subset of articles indexed in the Web of Science (i.e., 370 articles, 35 in national journals and 335 in international journals) published in 167 journals, the institutional IF per year, the mean institutional IF and the global value for InCor were calculated (Table 3). The InCor publications indexed from 2000 to 2003 generated

Table 1 - Articles published by authors from InCor, according to the databases in which they were indexed, from 2000 to 2003

Journal	Web of Science	MEDLINE	EMBASE	LILACS	SciELO	Not indexed	Total number of articles
National	35	133	126	244	181	1	246
International	335	353	340	0	0	3	358
<i>Total</i>	370	486	466	244	181	4	604
<i>% indexed</i>	61	81	77	40	30	-	99

Sources: databases and lists of journals indexed in 2005

Table 2 - Articles published by authors from InCor, according to the type of journal and mode of access, from 2000 to 2003

Mode of access	National journals		International journals	
	Total	%	Total	%
Electronic	223	91	346	97
Controlled access	2	1	333	93
Open access	221	90	13	4
Printed	23	9	12	3
Total	246	-	358	-

Sources: SciELO, Virtual Health Library, Capes Portal and the Internet, 2005

Table 3 - Number of articles published in journals indexed by the Web of Science, Institutional Impact Factors and means for the years 2000 to 2003

Year of publication	Number of articles	Institutional IF	Mean institutional IF
2000	77	165.5	2.149
2001	108	136.0	1.259
2002	76	343.5	4.520
2003	109	178.0	1.633
4-year period (2000-2003)	370	823.0	2.224

823 citations per year in the two subsequent years after their publication; moreover, the mean impact of each article by InCor was 2.224 in this period.

DISCUSSION

The assessment of the scientific production of academic institutions is an important measure of the extent of their contributions to developing new knowledge. Some indicators are traditionally used in this type of analysis, such as the number of publications, the indexing of journals where they were published and the number of times that an article has been cited by other publications.^{1,7,8,10}

InCor authors published 604 original articles and reviews during the 2000-2003 period, representing an average of approximately 150 publications per year. It is difficult to compare this finding with the production of other institutions.^{1,8} The result of such a comparison would be questionable since the publication habits differ significantly among scientific fields.

Of the 1253 retrieved records, 653 (52%) were letters, editorials, conference proceedings, case reports, brief communications or comments, demonstrating that these types of publications play an important role in communicating the

results of research carried out at the institution, despite their smaller relevance as compared to full articles and reviews.^{9,11} By analyzing records of the institutional database of the University Hospital of the Federal University of Rio de Janeiro, Araujo et al. found similar results. They classified publications other than original articles and reviews as “non-research-oriented articles”.⁸

The search strategy employed to retrieve the scientific production of InCor was based on combinations of the words that compose the name of the institution in the field of author affiliation. At this stage of the study, three limitations were observed concerning the search method by affiliation.^{2,7,9,12} The first refers to a lack of standardization by the institution for its identification in the articles submitted to publication – very often, the authors determine the nomenclature. The second concerns a lack of quality control of the journals that publish incomplete names of institutions. The third applies to varied standards adopted by the bibliographic databases as to the presentation of the author affiliation field. In the present study, these limitations were minimized by exploring terms that could possibly identify the InCor affiliation in each database and meticulously checking data from the retrieved articles against the institutional database.

Another finding that drew attention was that the existence of an institutional database contributed to a more complete retrieval of the scientific production of InCor and to an understanding of the search mechanisms in bibliographic databases. On the other hand, it was observed that approximately one third of the retrieved publications were not found in the InCor institutional database, suggesting that the current strategy to record and review the institute’s scientific production based on author’s notification of the InCor Scientific Documentation Service fails and institutional measures are necessary to improve this database.

No single database proved to be sufficient to retrieve the institutional scientific production using the above-mentioned search strategies. This finding is also true for other areas of knowledge, as Archambault et al. pointed out for the social sciences and humanities fields.⁴ Our results show that, for national journals, the database with the highest retrieval rate was LILACS (80%) and, for international journals, the Web of Science (90%). Based on these findings, we suggest that the retrieval of articles by Brazilian academic institutions should consider different bibliographic databases, both international and regional. For national journals, LILACS is recommended and, for international journals, the Web of Science and MEDLINE. Although the Web of Science showed higher retrieval rates than MEDLINE, the contents of both are complementary. EMBASE showed a high

proportion of overlap with MEDLINE in cardiology, which reduced its contribution to the results.

The preference for publishing articles in indexed journals, both in international and regional databases, indicated that InCor researchers are concerned with the visibility of their publications. The criteria adopted by the Capes to assess institutional production, which privileges indexed publications, may also have contributed to this trend.

The availability of 62% of international journals at the Capes Portal of Journals and of 74% of articles published in national journals in SciELO demonstrated the importance of these initiatives, which not only favor access to the international and national literature, respectively, but also influence the authors when they are selecting journals for publication.

Impact factors can be calculated in different ways.^{1,7,13} The method used in this study differs from the method proposed by Rousseau, for instance, in which the annual mean of citations over a three-year period is taken into account.¹ The proposed article IF presented in this study was intended to normalize the measure using a two-year period, making it comparable to the journal IF. Analysis of the impact of scientific production was limited to articles published in journals indexed by the Web of Science, due to the availability of indicators for IF and citations of articles in this database. This database proved to be an appropriate source for articles published in relevant international journals for this study.^{1,5} The proposed institutional IF has the advantage of being intuitive, easily comparable and potentially indicative of the quantity and visibility of the bibliographic production. The institutional IF per year is obtained by identifying the articles indexed in the Web of Science in a given year and calculating the sum of the article IFs. The mean of 2.224 per year presented by InCor during the period of study may be compared to the value of other institutions or be used for temporal review within the institution.

Another important institutional impact measure is

comparison of the total number of citations independent of the year of publication. This measure is not an appropriate comparison because it is influenced by the time since publication. In general, the number of citations increases as time goes by and, after reaching a maximum level, it starts decreasing.¹⁴

It is worth mentioning that there is significant variability in the number of citations of articles according to the field of knowledge and time since publication. Such variability has a significant effect on multidisciplinary institutions such as InCor, which engages in research in bioengineering, psychology, physiotherapy and molecular biology.

The quantitative analysis of scientific production is not sufficient to determine the quality and relevance of the scientific activities performed by research institutions. Other indicators, such as productivity of research groups, which can be measured by the time dedicated to research, collaboration with scientists of other groups and grant awards, would complement the assessment of institutional scientific productivity.¹⁻³ However, considering the difficulty of collecting detailed data on research groups, the strategy used in this study proved to be more feasible because it may be carried out using databases that are currently available on the internet.

The scarcity of studies with comparable data makes it difficult to contrast our results with those from other healthcare research institutions. Our approach suggests that a practical and simple methodology for bibliometric analysis of institutional scientific production consists of searches in the LILACS database for national journals and in the Web of Science and MEDLINE for international journals, and of calculating the institutional IF. InCor was assessed as an example and this study may represent a first step towards a broader understanding of the scientific production of healthcare research institutions when followed by studies in other institutions or by including further analyses exploring trends over time, types of publications or research fields and topics.

REFERENCES

1. Rousseau R. Indicadores bibliométricos e econométricos para a avaliação de instituições científicas. *Ci Inf Brasília*. 1998;27:149-58.
2. Moed HF, Burger WJM, Frankfort JG, Van Raan AFJ. The use of bibliometric data for the measurement of university research performance. *Res Policy*. 1985;14:131-49.
3. Pereira JCR, Pires ML, Duarte PS, Paes AT, Okano V. Introducing a method of research evaluation into a university: medical research at the University of São Paulo, Brazil. *Res Evaluation*. 1996;6:37-42.
4. Archambault E, Vignola-Gagne E, Côté G, Larivière V, Gingras Y. Benchmarking scientific output in the social sciences and humanities: the limits of the existing databases. *Scientometrics*. 2006;68:329-42.
5. Pereira JCR, Vasconcellos JP, Furusawa L, Barbati AM. Who's who and what's what in Brazilian Public Health Sciences. *Scientometrics*. 2007;73:37-52.

6. Zorzetto R, Razzouk D, Dubugras MTB, Gerolin J, Schor N, Guimarães JA, et al. The scientific production in health and biological sciences of the top 20 Brazilian universities. *Braz J Med Biol Res.* 2006;39:1513-20.
7. González-Sagrado M, de Luis Román DA, Conde-Vicente R, Izaola O, Aller R, Perez-Castrillón JL. Evaluación de dos métodos de corrección del Factor de Impacto utilizando la investigación del H.U. "Del Río Hortega" (1999-2004) como fuente de datos. *Nutr Hosp.* 2008;23:111-8.
8. Araújo KM, Mourão PAS, Leta J. Balance between education- and research-oriented publications from a Brazilian University Hospital. *Braz J Med Biol Res.* 2005;38:1285-91.
9. Montes GS. Distribution of financial resources according to productivity in the Hospital de Clínicas Medical Research Laboratories, University of São Paulo School of Medicine (Brazil). *Rev Med Chile.* 2000;128:431-6.
10. Kalternborn K-F, Kuhn K. The journal impact factor as a parameter for the evaluation of researchers and research. *Rev Esp Enferm Dig.* 2004;96:460-76.
11. Petroianu A. Critérios quantitativos para analisar o valor da publicação de artigos científicos. *Rev Assoc Med Bras.* 2003;49:173-6.
12. Meneghini R. Systematization of academic and scientific affiliation, or how to prevent data on your publications from being lost in the national and international data base. *Braz J Med Biol Res.* 1995;28:617-9.
13. Habibzadeh F, Yadollahie M. Journal weighted impact factor: A proposal. *J Informetrics.* 2008;2:164-72.
14. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A.* 2005;102:16569-72.