Mapping of breast cancer research in India: a bibliometric analysis

Nirmal Singh*, Tarvinder Singh Handa, Dhiraj Kumar and Gurpreet Singh

This study presents a bibliometric analysis of the scholastic output on breast cancer in India. The purpose is to provide an overview of the research activities in the country on the subject during the last ten years, exploring different aspects of scientific literature. Data on 3529 items, including 2945 articles and 584 reviews published during 1 January 2005 to 31 December 2014 were collected using Scopus. An advanced search was conducted in the database. The search results were filtered for English language, journals, India and the period under study. Lotka’s law was applied to assess the author productivity and Bradford’s law of scattering was used to ascertain the distribution pattern of articles in journals. Most (about 96%) contributions were found to be an outcome of collaborative authorship. Around 19.05% of the papers had collaboration of four authors and 16.33% appeared due to collective efforts of three authors each. Trend of an increasing number of articles published over the period has been observed. Author productivity did not fit the Lotka’s law with a value of \( n = 2 \). The distribution of articles in journals was found acceptable to the Bradford’s law of scattering.

Keywords: Bibliometric analysis, breast cancer, collaborative authorship, research activities.

Cancer is one of the major health concerns worldwide. The year 2012 witnessed 14.1 million new cases of cancer and global mortality of 8.2 million due to the noxious disease. Cancer of the lungs, colorectal and stomach cancer and breast cancer accounted for more than 40% of all cases diagnosed globally. While cancer of the lungs occurred more frequently in men, breast cancer was the most commonly diagnosed cancer in women, constituting 25.2% of all new cases.

Cancer is one of the top 10 causes of death in India. Breast cancer accounts for 22.2% of all new cancer diagnoses and 17.2% of all cancer deaths among women in the country. In 2012, 144,937 women were newly detected with breast cancer and 70,218 died due to this fatal disease. The rate of increase in breast cancer in the country is so rampant, that if we do not act now, we are in for a major shock in the next twenty years. More alarming is a considerable shift over the last few decades in the age of young women showing development of breast cancer. While the peak occurrence of breast cancer in the United States and other European countries in women is in their sixties, almost 48% of the patients in India are below 50, with an increasing number of patients in the age group between 25 and 40 years.

The world age-standardized rate (ASR) of incidences of breast cancer is 43.1 per 100,000 women and the age-standardized breast cancer death rate is 12.9 per 100,000 females. Compared to this, the ASR of incidences of breast cancer in India is 25.76 per 100,000 women. However, in proportion to the incidence rate, the ASR of mortality in the country is much higher than that in the developed countries, taking 12.73 lives per 100,000 women. Hence, in India, on an average, for every two women newly diagnosed with breast cancer, one is dying from this disease. It is estimated that the incidence of new cases of breast cancer in India will rise up to 200,000 per year by 2030 (ref. 11).

Women have a crucial role in the social, economic and cultural development of society. In male-dominated India, the health and care of women is a critical issue. The absence of typical female advantage in life expectancy suggests the systematic problems in the health of women in the nation. Women are mainly exposed to indoor pollutants at home and in the workplace, and there are evidences that they are more vulnerable than men to various chemicals. Improper nutritional intake, overweight and obesity in early adulthood also account for this fatal disease. For sustainable well-being of women, it is essential that strategic interventions are made at critical stages. Research, evidence and information are basic to sound health policies. There is significant positive

Nirmal Singh and Dhiraj Kumar are in the Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana 141 004, India; Tarvinder Singh Handa is in the Indian Institute of Technology Ropar, Rupnagar 140 001, India; Gurpreet Singh is in the Goswami Ganesh Dutt College, Chandigarh 160 031, India.

*For correspondence. (e-mail: nirmal02@yahoo.co.in)
correlation between improvement in the survival rate of patients and research output. Therefore, quantitative assessment of the scholarly research output of India on breast cancer is pertinent to map out the growth trends and its future perspectives, as this can have a significant bearing on the future research and policies to tackle the disease in gender-disturbed nation. Bibliometric analysis is a statistical support device to map out and generate different types of information, and knowledge handling and management indicators.

**Objectives**

The present study is aimed at assessing (i) the year-wise distribution of papers, (ii) authorship pattern, (iii) author productivity and productive authors, (iv) institutional contributions, (v) fitness of distribution of papers to Bradford’s law of scattering, and (vi) the core journals publishing papers on breast cancer.

**Methodology**

For bibliometric analysis, data regarding papers on breast cancer were collected using Scopus in July 2015, for the period 1 January 2005 to 31 December 2014. Scopus is the largest abstract and citation database of peer-reviewed literature, including scientific journals, books and conference proceedings in the disciplines of science, technology, medicine, social sciences, arts and humanities. To access data for the present study, Scopus was searched as follows:

- An advanced search was conducted using ‘breast cancer’ and alternative search terms, i.e. ‘breast carcinoma’, ‘neoplasm of breast’, ‘tumor of breast’, ‘tumour of breast’, ‘mammary cancer’, ‘ductal carcinoma’ and ‘invasive carcinoma’.
- The search was restricted to the occurrence of search terms in the title, abstract and keywords of the articles.
- Boolean operator ‘OR’ was applied to the above search terms to produce an exhaustive number of results.
- The search results were filtered to English language, journals, India and restricted to the period of ten years under study, i.e. 2005–2014.
- The search was further filtered by subject areas to cover articles in medicine, biochemistry, pharmacology, nursing, health science, multidisciplinary and immunology.
- The search results provided 3529 records covering 2945 articles and 584 reviews for the period under study on the given search terms.
- The data were downloaded in Excel format and analysed.

- Data analysis was performed using frequencies and percentages of publications. Besides, Lotka’s law was applied to assess the author productivity and Bradford’s law of scattering was used to ascertain distribution pattern of articles.

**Analysis and discussion**

**Year-wise distribution of papers**

During the period 2005–2014, a total 3529 papers were published on breast cancer by authors with institutional affiliation in India, either individually or in national/international collaboration. The trend of growing number of publications over the period can be observed in Table 1, indicating that with the increasing burden of breast cancer in the country, research on the issue has also increased. This supports the study of Kotepui et al.20 revealing the trend of growing number of publications from Asian countries. It took seven years (2005–2011) to produce 48.54% (1713 papers) of the total contributions during the period under study, while the rest 51.46% was published during the last 3 years (2012–2014) only. The year 2014 produced more publications than the first four years under study, viz. 2005–2008. On an average,
scientific output on the subject increased by nearly 19% per annum.

Authorship pattern

The number of authors contributing to each publication varied from 1 to 295. A large majority of papers had been written in collaboration, with the exception of only 135 (03.82%) single-authored papers (Table 2), corroborating the findings of Minas et al. and Sridevi. Interestingly, 05.41% of the papers resulted from collaboration of more than 10 authors, including 4 articles having more than 100 authors each. Also, 3529 articles and reviews were contributed by 18,544 authors, each publication having around 5 authors, on an average. The degree of collaboration of authorship was calculated using the formula given by Subramaniam:

\[ C = \frac{Nm}{Nm + Ns} = \frac{3394}{3394 + 135} = 0.9617, \]

where C is the degree of collaboration, Nm the number of multi-authored works, and Ns is the number of single-authored works.

Author productivity

Author productivity was assessed considering the first author of each article (Table 3). A total of 2521 authors had made 3529 contributions. Majority of authors (80.00%) contributed only one paper, followed by 11.79% authors contributing 2 publications each. Lotka’s law was applied to calculate the number of expected authors for a given number of publications. Considering the fact that 2017 authors have produced only 1 paper each, the value of \( n \) can easily be derived. Putting the value of \( n \) as 2, the results shown in Table 3 were obtained.

Table 3 shows that there are only a few productive authors, a large majority contributing to breast cancer occasionally. In contrast to the study of Parta and Bhattacharya, the present results suggest that in this case the author distribution does not obey Lotka’s law. The difference between the number of observed authors and expected authors was considerably wider.

Prolific authors

The most prolific authors on the subject of breast cancer have been identified and ranked (Table 4). There were 12 authors each publishing more than 20 papers. The author with the highest contribution in terms of the number of papers on breast cancer is Sachdanandam contributing 30 publications, followed by Parshad and Konwar (26 each). Saxena and Sarin contributed 25 and 24 papers respectively. Of the total 18,544 authors, 159 Indian authors made a contribution to 51.40% of the total 3529 publications, whereas the remaining majority to the rest of the 48.60% contributions.

Core journals in the subject

The total 3529 papers analysed in this study appeared in 972 journals from various publishers and geographical locations. Table 5 shows the most productive journals on breast cancer. Around 27.76% (980) of the total publications under the study appeared only in 25 journals, which may be considered as core journals (Table 5).

The impact factor (IF) is a widely accepted quality determinant for journals, reflecting the average number of citations to recent articles published in a journal. Higher the IF, more important the journal is considered to be. However, 10 of the 25 core journals publishing the highest number of articles on breast cancer did not have an IF. Eight of these 10 journals without IF, are being...
published from India. This corroborates the findings of Patra and Bhattacharya, that the Indian cancer research articles are not published in high-impact journals.

Bradford’s law of scattering

Bradford’s law of scattering is used to describe the distribution of the literature on a particular subject in journals. The law working on mathematical means is based on the principle of centric productivity zones, demonstrating that there are diminishing returns when the literature is published exhaustively. According to the law, journals can be divided into different zones containing the same number of articles. For example, the core zone contains one-third of the total articles; similarly zone 1 contains the same number of articles, but a greater number of journals, and zone 2 contains the same number of articles, but still greater number of journals, and so on. This increase in the number of journals from one zone to the next is according to the expression $1 : n : n^2$. The law helps to distinguish the groups of journals dedicated more specifically to the subject of interest. The number of journals in each zone can be calculated from Bradford’s multiplier constant $k$. In this study, $k$ has been arrived at using the mathematical formulations of Egghe and Rousseau as given below

$$k = (e^y \times Ym)^{1/p},$$

where $y$ is Euler’s number having a value 0.57772, $Ym$ is the number of articles published in the top-ranked journals and $p$ is the Bradford group or number of zones, i.e. $p = 3$. Therefore,

$$k = (1.781 \times 104)^{1/3} = 5.70.$$  

The different Bradford groups can be calculated using $k$. The core zone $r_0$ can be defined as

$$r_0 = \frac{T(k-1)}{(k^p-1)},$$

where $T$ represents the total number of journals in the study. Thus,

$$r_0 = \frac{972(5.70-1)}{(5.70^3-1)} = \frac{4568.4}{184.193} = 24.81.$$
be observed that medical institutions and hospitals make a greater contribution to the literature on breast cancer than universities, contrary to the findings of Ortiz et al.28.

Collaboration with other nations

Table 8 lists the major nations along with the number of papers in which Indian authors had collaborated with the authors from other nations. Indian authors contributed 417 publications in collaboration with authors from the United States. India has collaboration in 57 papers with Germany, followed by 56 contributions with the United Kingdom, Canada, France, Saudi Arabia, Australia, Singapore and South Korea are the other nations with which Indian authors have contributed 30 or more papers.

Conclusion

Research and scientific activities on breast cancer involve a high degree of collaboration, not limiting to the geographical boundaries. The present study shows that 11.81% of the papers are contributed by Indian authors in collaboration with authors from the United States. Similarly, countries such as Germany, United Kingdom, Canada, France, etc. collaborate with authors from India. A growing trend of publications in the subject is observed. In comparison to 119 papers published in 2005, the number of contributions has increased by 466% during 2014. Majority of the authors contribute to the subject occasionally. However, author productivity does not fit to Lotka’s law, as the observed and expected values vary significantly. The distribution of papers obeys the Bradford’s law of scattering identifying 25 core journals. The

<table>
<thead>
<tr>
<th>Institution</th>
<th>No. of papers</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tata Memorial Hospital</td>
<td>175</td>
<td>04.95</td>
</tr>
<tr>
<td>All India Institute of Medical Sciences</td>
<td>169</td>
<td>04.78</td>
</tr>
<tr>
<td>University of Madras</td>
<td>90</td>
<td>02.55</td>
</tr>
<tr>
<td>Indian Institute of Chemical Technology</td>
<td>81</td>
<td>02.29</td>
</tr>
<tr>
<td>Postgraduate Institute of Medical Education and Research</td>
<td>77</td>
<td>02.18</td>
</tr>
<tr>
<td>Central Drug Research Institute</td>
<td>63</td>
<td>01.78</td>
</tr>
<tr>
<td>National Institute of Pharmaceutical Education and Research</td>
<td>54</td>
<td>01.53</td>
</tr>
<tr>
<td>Sanjay Gandhi Postgraduate Institute of Medical Sciences</td>
<td>47</td>
<td>01.33</td>
</tr>
<tr>
<td>Chittaranjan National Cancer Institute</td>
<td>46</td>
<td>01.30</td>
</tr>
<tr>
<td>National Centre for Cell Science</td>
<td>42</td>
<td>01.19</td>
</tr>
<tr>
<td>Indian Institute of Science</td>
<td>41</td>
<td>01.16</td>
</tr>
<tr>
<td>Nizam's Institute of Medical Sciences</td>
<td>40</td>
<td>01.13</td>
</tr>
<tr>
<td>Banaras Hindu University Institute of Medical Sciences</td>
<td>40</td>
<td>01.13</td>
</tr>
<tr>
<td>Osmania University</td>
<td>39</td>
<td>01.10</td>
</tr>
<tr>
<td>Institute Rotary Cancer Hospital</td>
<td>38</td>
<td>01.07</td>
</tr>
<tr>
<td>Annamalai University</td>
<td>37</td>
<td>01.04</td>
</tr>
<tr>
<td>Regional Cancer Centre</td>
<td>36</td>
<td>01.02</td>
</tr>
<tr>
<td>Vardhman Mahavir Medical College and Safdarjung Hospital</td>
<td>36</td>
<td>01.02</td>
</tr>
<tr>
<td>Manipal University</td>
<td>36</td>
<td>01.02</td>
</tr>
<tr>
<td>Jadavpur University</td>
<td>36</td>
<td>01.02</td>
</tr>
</tbody>
</table>

**Table 7. Institutional contribution**

Different Bradford zones have been obtained using the values of $k$ and $r_0$:

- Core zone $r_0 = r_0 \times 1 = 24.81$,
- First zone $r_1 = r_0 \times k = 24.81 \times 5.70 = 141.42$,
- Second zone $r_2 = r_0 \times k^2 = 24.81 \times 5.70^2 = 806.08$.

The above theoretical distribution according to Bradford’s law enables one to examine the exact fit of the law to the distribution of articles. Using this distribution, the number of journals in each zone has been arrived at Table 6. Using the distribution of journals in Table 6, $k$ is 5.700 and 5.699 for zone 1 and zone 2 respectively. This value of $k$ is similar to that calculated using the formula $k = (c' \times Y m)^{1/p}$. This makes it clear that data collected for the present study fit into the three zones of Bradford’s law of scattering, i.e. $1 : k : k^2$ or $1 : n : n^2$.

**Institutional contribution**

Table 7 gives the top 20 institutions in terms of their contribution to the number of papers on breast cancer. It can

<table>
<thead>
<tr>
<th>Country</th>
<th>No. of papers</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>3529</td>
<td>100.00</td>
</tr>
<tr>
<td>United States of America</td>
<td>417</td>
<td>11.81</td>
</tr>
<tr>
<td>Germany</td>
<td>57</td>
<td>01.61</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>56</td>
<td>01.58</td>
</tr>
<tr>
<td>Canada</td>
<td>48</td>
<td>01.36</td>
</tr>
<tr>
<td>France</td>
<td>45</td>
<td>01.27</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>43</td>
<td>01.21</td>
</tr>
<tr>
<td>Australia</td>
<td>41</td>
<td>01.16</td>
</tr>
<tr>
<td>Singapore</td>
<td>32</td>
<td>00.90</td>
</tr>
<tr>
<td>South Korea</td>
<td>30</td>
<td>00.85</td>
</tr>
</tbody>
</table>

**Table 8. Collaboration with other nations**
inputs to the scientific literature are dominated by a few selected institutions. It is pertinent to mention here that the present study is limited to bibliometric analysis of the scientific literature on breast cancer contributed by India. A comparative study with scientific inputs from other nations on the subject will provide a picture of research trends in the global context.


Received 12 August 2015; accepted 19 October 2015
doi: 10.18520/cs/v110/i7/1178-1183