Croatian Bibliometric Analysis, 2000–2007

Dan S. Radut and Javier Sanz-Valero

1 University of Alicante, Department of Community Nursing, Preventive Medicine, Public Health and History of Science, Spain
2 Miguel Hernandez University, Department of Public Health, History of Science and Gynecology, Elche, Spain

ABSTRACT

To develop search filters and retrieve information estimating the Croatian scientific output (SO) focusing on Public Health (PH) and Preventive Medicine (PM) in MEDLINE. A PubMed search of the MEDLINE database was performed to retrieve articles added to this database between 2000 and 2007. Search filters inspired by previous strategies were applied involving ‘geographical’, ‘place of publication’, ‘subject’ and ‘language of publication’ aspects. An evaluation of the geographical filter performance was done and sensitivity and specificity were calculated. There were obtained publications in several languages, originated in Croatia, published in Croatia and/or abroad. The Croatian SO in the field of PH-PM was obtained for the same period of time by combining search filters. The evaluation of the filter performance showed sensitivity 95.56% and specificity 100%. The filters constructed permitted the retrieval of the Croatian eight years research output. Increased tendency was observed in the global SO evolution and in the PH-PM area as well. The main languages of publication were English and Croatian. This study is a contribution to research in the field of scientific documentation and further analysis is recommended in constructing and developing search filters to retrieve and focus on specific information.

Key words: bibliometrics, databases, bibliographic, public health, preventive medicine, Croatia

Introduction

Scientific journals play an important role in the transfer of scientific information, and international visibility is a prerequisite for fulfilling this role1. In the information era, every year only the electronic database Medline reports 560,000 new scientific articles. In such a context the search filters become useful and recommended tools for researchers to determine specific information. The search filters’ validation depends on the availability of a ‘gold standard’ – a reference standard against which to establish the performance of the filter2,3. Search filters encompass different aspects such as geographical, to help finding information corresponding to specific location, or thematic focusing on specific field of study. Methodological search filters are defined by a search strategy based on a combination of terms related to each other and which enable the retrieval of specific information. In other words, these filters are tools whose main objectives are to attenuate the negative effects caused by the excess of information in the bibliographic databases, to reduce instances of superfluous information in search results, as well as to maximize the specificity and sensitivity of the search4. Researchers and clinicians are increasingly utilizing search filters in their literature searches but the lack of literature examining search filter methodology has been noted5.

The bibliometric studies can be used to determine and/or analyze the amount of research product – scientific output (SO) – in national/international context, fields of activity and determined period of time. In analyzing the SO, the bibliometric indicators are useful tools, provided their interpretation is not extrapolated beyond their limits of validity6.

Analyzing the research product in the field of Public Health and Preventive medicine becomes even more important since the research performed in these areas provides the basis for identifying significant health problems in the population and supports the development of programmes to educate people on health issues, promote health and protect people’s vulnerability to different health hazards. In the field of Public Health (PH) and Preventive Medicine (PM), some data reveal that between 1995 and 2003, Europe produced 27% of the Public Health worldwide research production7 but 26.3% was...
originated from Western Europe and 0.7% from Eastern Europe.

The aim of this study was to develop and apply a generic search strategy for the retrieval of the Croatian SO’s characteristics (both globally and in the field of PH and PM) recorded in MEDLINE (accessed through PubMed) between 2000 and 2007.

Materials and Methods

A bibliographic search has been carried out on a year-by-year basis between 2000 and 2007, by using search filters to find the articles in any language originated from Croatia and added to MEDLINE database. Data was extracted from MEDLINE through PubMed, the leading online public database of biomedical literature records and an essential tool used by scientists and physicians to monitor research developments in any given field of science or medicine8. The search was carried out on December 17th 2008, the frequencies and percentages were determined. The search filters were constructed by using Medical Subject Heading (MeSH) Terms and Tags according to affiliation (geographical filters), field of study (thematic filters) and place of publication. The Boolean operators ‘OR’, ‘AND’ and ‘NOT’ were used to cross/combine the equations.

Constructing geographical search filter

A geographical filter was constructed using the subheading ‘Croatia’ with the affiliation Tag ‘[ad]’ added. By using the Boolean ‘OR’ and despite that the results should be encompassed by the simple term ‘Croatia’ for more certainty the following alternative it was also included in the filter: ‘Republic of Croatia’. The translation of these terms was done in Croatian and some of the main languages used in the scientific journals’ field, including: German, French, Italian, Spanish and Portuguese. All the terms were combined by using the Boolean ‘OR’. Other parameters taken into consideration when constructing the filter were the Croatian counties and main cities as follows: Zagreb, Krapina, Krapina Zagorje, Sisak, Sisak Moslavina, Karlovac, Varazdin, Koprivnica, Koprivnica Krizevci, Bjelovar, Bjelovar Bilogora, Rijeka, Primorje Gorski Kotar, Gospic, Novalja, Otokac, Senj, Lika Senj, Virovitica, Virovitica Podravina, Pozega, Pozega Slavonia, Slavonski Brod, Brod Posavina, Zadar, Osijek, Osijek Baranja, Sibenik, Sibenik Knin, Vukovar, Vukovar Srijem, Split, Split Dalmatia, Pazin, Pula, Istria, Dubrovnik, Dubrovnik Neretva, Cakovec and Medimurje. The terms were added to the filter in both languages, English and Croatian. The use of geographical search filters was analyzed in detail by Valderas et al.9.

In order to obtain the thematic search filter, the Medical Subject Heading (MeSH) Database was accessed and the Medical Subject Heading Terms ‘Public Health’ and ‘Preventive Medicine’ were selected and combined in a common search equation by using the Boolean ‘OR’. The equation obtained was added to the geographical filter with the Boolean ‘AND’.

To detect the articles published in Croatia and/or the articles published abroad, the name of the country was added to the ‘place of publication’ Tag ‘[pl]’. This equation was then crossed with the ‘geographical’ and ‘field of study’ filters by means of the Boolean ‘AND’. The articles were distributed into two categories: ‘Published in Croatia’ and ‘Published abroad’ (articles not published in Croatia).

In all cases, the search was carried out for English and/or Croatian language by using the language option permitted by the database search system. To detect bilingual publications, the language Tag ‘[la]’ was added to the languages which were then combined with the Boolean ‘AND’ and added to the rest of the filter with the Boolean ‘AND’ [i.e. (English[la] AND Croatian[la]) AND]. To detect other languages, English/Croatian were excluded with the Boolean ‘NOT’ and then by using a manual revision of the publications. The articles were categorized according to ‘Croatian’, ‘English’ and the references in other languages were classified into ‘other language’.

Evaluation of geographic filter performance

The geographic filter performance was evaluated comparing the results obtained by means of a manual revision to the results obtained after the application of the geographical filter.

The identification of the sample and the manual search

The sample to evaluate consisted of all articles published in the area of PH-PM in Croatia and added to PubMed (Medline) between 2000 and 2007. The manual revision then classified the articles into ‘relevant’ and ‘not relevant’. Relevant articles were those that specified the study was carried out in Croatia, written in any language by any author belonging to any Croatian research institution. The articles which did not fulfill these conditions were categorized as ‘not relevant’. The results were compared to the results obtained after the application of the geographical filter on the base sample, and analyzed by following the gold standard reference set model10. Sensitivity and specificity were calculated. The statistical analysis was performed with the SPSS 11 program. Frequencies and percentages were determined.

Results

At the time of the search, and according to the methodology chosen, 6,079 publications were found dated between 2000 and 2007 and corresponding to the Croatian SO. The annual distribution and the characteristics regarding the language and place of publication are shown in Table 1. According to the language of publication, four bilingual articles English-Croatian were found corresponding to 2006 (two) and 2007 (two). These articles were counted for SO but were not included in either English or Croatian language category articles. Excluding Croatian and English, 46 articles were written in other...
languages: two (published in Croatia) in Czech in 2000, three in German and one (published in Croatia) in Serbian in 2001, three in German in 2002, fifteen (published in Croatia) in Polish and three in German in 2003, six in German, one in French, one in Italian and one in Bosnian in 2004, two in German and one in Bosnian in 2005, two in German, one in Italian and one in Bosnian in 2006 and three in German in 2007 (Table 1).

2,425 (39.89%) articles of the total Croatian SO found were published in Croatia as follows: 188 (7.75%) articles in 2000; 172 (7.09%) in 2001; 264 (10.89%) in 2002; 331 (13.65%) in 2003; 277 (11.42%) in 2004; 398 (16.41%) in 2005; 369 (15.22%) in 2006 and 426 (17.57%) in 2007. Almost all of the articles were published in English, except for the following: one in Croatian (published in Bosnia) in 2000; three in German in 2001; eight in Croatian (published in Bosnia) and three in German in 2002; four in Croatian (published in Bosnia) and three in German in 2003; six in German, one in French, one in Italian and one in Bosnian in 2004; two in German and one in Bosnian in 2005; two in German, one in Bosnian and one in Italian in 2006, and three in German in 2007.

The publications in this field were retrieved applying the PH and PM filter. A total number of 2,458 (40.43%) publications were obtained. 1,404 (57.12%) of these articles were published abroad (Table 1). In Croatia, mainly in the journals 'Collegium antropologicum', 'Croatian Medical Journal' and 'Acta medica Croatica' (approximately 80% of the SO published in this country) were detected 1,054 (42.88%) publications as follows: 89 (8.44%) in 2000; 74 (7.02%) in 2001; 123 (11.67%) in 2002; 157 (14.90%) in 2003; 99 (9.39%) in 2004; 171 (16.22%) in 2005, 154 (14.61%) in 2006 and 187 (17.74%) in 2007. Excluding four articles in Polish in 2003, all of the articles belonging to the fields of PH or PM and published in Croatia were in English or Croatian. With regard to the PH-PM articles published abroad, excluding one article in Croatian in 2000, two in German in 2001, two in German and one in Croatian in 2002, two in German and one in Croatian in 2003, five in German in 2004, two in German in 2005 and one in German in 2007, all publications were written in English.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO</td>
<td>494</td>
<td>544</td>
<td>644</td>
<td>734</td>
<td>730</td>
<td>946</td>
<td>944</td>
<td>1043</td>
<td>6079</td>
</tr>
<tr>
<td>Language</td>
<td>Cro.</td>
<td>37</td>
<td>60</td>
<td>71</td>
<td>119</td>
<td>125</td>
<td>142</td>
<td>154</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>Engl.</td>
<td>455</td>
<td>480</td>
<td>570</td>
<td>597</td>
<td>596</td>
<td>801</td>
<td>784</td>
<td>5164</td>
</tr>
<tr>
<td>Published in Croatia</td>
<td>36</td>
<td>60</td>
<td>63</td>
<td>115</td>
<td>125</td>
<td>142</td>
<td>154</td>
<td>182</td>
<td>1544</td>
</tr>
<tr>
<td>Published abroad</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
<td>2007</td>
<td>3654</td>
</tr>
</tbody>
</table>

TABLE 1
CROATIAN ANNUAL DISTRIBUTION OF ARTICLES BY SO, LANGUAGE AND PLACE OF PUBLICATION IN NUMERICAL AND (PERCENTAGE TERM)

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO: PH-PM</td>
<td>198</td>
<td>221</td>
<td>270</td>
<td>299</td>
<td>299</td>
<td>399</td>
<td>399</td>
<td>426</td>
<td>2458</td>
</tr>
<tr>
<td>Language</td>
<td>Cro.</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>35</td>
<td>25</td>
<td>28</td>
<td>37</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>Engl.</td>
<td>191</td>
<td>210</td>
<td>256</td>
<td>258</td>
<td>247</td>
<td>369</td>
<td>331</td>
<td>2242</td>
</tr>
<tr>
<td>Published in Croatia</td>
<td>6</td>
<td>9</td>
<td>11</td>
<td>34</td>
<td>25</td>
<td>28</td>
<td>37</td>
<td>45</td>
<td>195</td>
</tr>
<tr>
<td>Published abroad</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>18</td>
</tr>
</tbody>
</table>

TABLE 2
CROATIAN ANNUAL DISTRIBUTION OF ARTICLES BY SO, LANGUAGE AND PLACE OF PUBLICATION IN NUMERICAL AND (PERCENTAGE TERM) IN THE FIELD OF PUBLIC HEALTH (PH) AND PREVENTIVE MEDICINE (PM)

1 Two articles posted in both English and Croatian in 2006 were added to SO category but not to language individual category
2 Two articles posted in both English and Croatian in 2007 were added to SO category but not to language individual category

The affiliation of the Croatian research centers originating publications in the field PH-PM was checked through a manual revision of all the articles in this area. The revision confirmed correct inclusion for a number of 2,451 (99.72%) articles and incorrect inclusion for 7 (0.28%) publications. The incorrect retrieval was owed to the following causes: 3 cases because the similar names of two cities located in Croatia and Italy, 'Pula'; 2 cases owed to the similar names of the Croatian city Split and a part of the electronic mail used by the authors of the articles; 1 case due to the similar names of the Croatian city name 'Gospic' and the family name of an author and 1 case because the similar names of a research author group of an article and the Croatian city Split.

The evaluation of the filter’s performance was carried out on a sample of 1,783 references, encompassing the articles belonging to the field of PH and PM, published in Croatia and added to MEDLINE between 2000 and 2007. The manual revision classified a number of 1,103 (61.86%) references as relevant and 680 (38.14%) as not relevant. By applying the geographical filter 1,054 publications were retrieved, all of them being found as relevant by means of the manual revision. The filter did not retrieve 49 references considered relevant by the manual revision. On the whole, the filter provided 95.56% sensitivity and 100% specificity (Table 3).

### Discussion

With the purpose of determining the Croatian SO with its dynamics on an eight year basis and focusing on the field of PH and PM in MEDLINE, a search process was carried out by developing search filters for this database. The geographical filter enabled the retrieval of scientific publications originating in Croatia, and information belonging to the area of PH and PM was retrieved by using a thematic filter. The 'place of publication' filter enabled the differentiation between references published in Croatia and abroad.

According to Misak et al.11 in the year 2000 six Croatian journals were indexed for MEDLINE; almost double that number of journals is indexed in 2008 (Table 4). The journals indexed in 2008 were gathered from the List of Journals Indexed (LJI) for MEDLINE12 and the additional data about the journals was obtained by searching for each journal individually in the section 'journals' in PubMed according to the journal’s name. In 2000 three journals published articles in English, two journals in Croatian and one journal was bilingual (English/Croatian). Currently, two journals indexed for MEDLINE publish in English, three in Croatian, four in English/Croatian. One journal receives articles in either English or German and one journal in English or multiple languages. The results of this study show the Croatian research centers' preference for English in both categories of articles published in Croatia or abroad. The dynamic of the SO published in Croatia related to the language of publication (Figure 1) shows quite similar evolution in both cases, global and PH-PM SO. More articles were published abroad than in Croatia and in both cases, published in Croatia or abroad, every year English was more used than Croatian. Nevertheless, as a positive fact it should be pointed out that the articles in the Croatian language also followed an increasing evolution in both global and/or PH-PM fields, especially between 2002 and 2003 when an important rise is observed (Figure 1). The
articles published abroad were written in English almost in totality and the main non-English languages used were German and Croatian (Figure 2). The international visibility is important in the field of research and we consider positive the extensive use of English providing an admirable ‘internationalization’ of the Croatian research; however this could present a disadvantage, the English publications in Croatia limiting access to only those who understand this language.

Similar bibliometric studies were carried out by Soteriades y Falagas7, who reveal an international map of the PH and PM research area, focusing also on subcategories like epidemiology, occupational medicine and environmental medicine. In the field of bibliometry, Valderas et al.9 described the geographical filters and Fraser et al.13 described the use of the methodological search filters. Also analysed were the bibliometric aspects with their characteristics in MEDLINE14 and Polasek et al.15 related the SO to the academic field. Other bibliometric studies identified SO in molecular biology, urology or dermatology16–18.

The application of the high sensitive (95.56%) filter presented in this article permitted the determination of the publications originating in Croatia and the thematic filter permitted to focus on the fields of PH and PM.

The geographical filters based on the use of affiliation Tags in PubMed do not detect articles without affiliation section. This fact may represent a limitation with regard to the obtaining of greater information when searching MEDLINE database through PubMed by using geographical filters like those presented in this study.

In summary, by using search filters we obtained a ‘radiography’ of the Croatian research in MEDLINE in the last eight years (except 2008) with respect to the global Croatian scientific output and the field of PH and PM, observing the dynamic of the publication process and focusing on parameters related to the languages used and also the place of publication (national or abroad). The type of filters presented in this can have great utility with regard to the bibliometric analysis and strategies and in finding specific information for researchers. The results of this article are tentative and considering this
study as a contribution to research in the field of scientific documentation, complementary studies are recommended to focus on the improvement of geographical search filters as presented in this article. We consider that further analysis is needed to advance in developing search strategies in order to detect, analyze and find precise information.

Acknowledgements

Our special thanks go to Prof. Antonio Garcia Belmar, Department of Community Nursing, Preventive Medicine, Public Health and History of Science, University of Alicante, for his excellent academic support, encouragement and timely review and comments about the article.

REFERENCES


D. S. Radut

University of Alicante, Department of Public Health, C/San Vicente del Raspeig s/n 03690 Alicante, Spain
e-mail: radut@ua.es

HRVATSKA BIBLIOMETRIJSKA ANALIZA, 2000–2007

S A Ž E T A K

Cilj je bio razviti filtere pretrage i pronaći informacije koje ocijenjuju Hrvatski znanstveni ispis s fokusom na javno zdravstvo i preventivnu medicinu u MEDLINE-u. PubMed pretraživanje MEDLINE baze informacija napravljeno je tako da se pronadu radovi dodani u ovu bazu između 2000. i 2007. godine. Filteri pretrage uključivale su geografski aspekt, aspekt mjesta objavljivanja, subjekta te jezika publikacije. U upotrebi su bile i publikacije na nekoliko jezika, izvorno iz Hrvatske, publicirane u Hrvatskoj i/ili u inozemstvu. Hrvatski znanstveni ispis na području javnog zdravstva i preventivne medicinе napravljen je za navedeni period s kombinirajućim filterima. Vrednovanje osjetljivosti učinka filtra pokazuje osjetljivost od 95,56% i specifičnost od 100%. Vodeći jezik publikacija bio je engleski i hrvatski. Ova studija doprinos je istraživanju znanstvene dokumentacije, a buduće analize poljene su u području konstrukcije i razvoja filtra za traženje kako bi se pronašlo i fokusirano na specifične informacije.