Pergamon

European Journal of Cancer, Vol. 35, No. 8, pp. 1182–1186, 1999

Printed in Great Britain
0959-8049/99/$ - see front matter

PII: S0959-8049(99)00107-0

Special Paper

Impact Assessment of Oncology Research in the European Union

G.S. Mela,[1] M.A. Cimmino[1] and D. Ugolini[2]

[1]Dipartimento di Medicina Interna e Specialita' Mediche, Universita' di Genova; and [2]Dipartimento di Oncologia Clinica e Sperimentale, Universita' di Genova, Centro Documentazione, Largo R. Benzi, 10, 16132 Genova, Italy

In this study the distribution of papers published by authors from the European Union (EU) in oncological journals was analysed, as was the impact of oncological research in the EU compared with that produced in other countries. Papers published during 1995 in the oncological journals listed by ISI (Institute for Scientific Information, Philadelphia, U.S.A.) were downloaded. The parameters of impact factor (IF), source country population and gross domestic product (GDP) were considered. An analysis of the key words, both those reported by the authors and those attributed by ISI, was carried out using a special purpose program. 36.5% of papers published in oncological journals come from the EU (the U.K., Italy, Germany and France ranking at the top) and 40.7% from the U.S.A. The mean IF was 2.4 for EU papers, 3.3 for the US and 2.4 for other countries. Our data confirm that smaller countries performed better than larger ones. The key words analysis shows that the leading fields of research were breast cancer for diseases, cisplatin for drugs and p53 for experimental studies. A standardisation of key words on behalf of journal editors is proposed.

Key words: neoplasms, bibliometrics, Europe
Eur J Cancer, Vol. 35, No. 8, pp. 1182–1186, 1999

INTRODUCTION

In recent years, economic difficulties due to growing national debts have increasingly forced governments, normally the primary supporter of basic research, to adopt policies that link science and technology programmes more closely to broad organisational and societal goals.

Consequently, the assessment of research output has progressively developed and necessarily become a priority issue for the scientific research community. However, quantifying and weighting the results of research are difficult and debated tasks. Several studies on the subject have shown the potential validity of particular bibliometric measures, and bibliometric analysis is largely used as an assessment tool of scientific research performance at different levels, be it a discipline oriented [1–3], individual [4, 5], national [6, 7], regional or global [8] evaluation process.

Correspondence to D. Ugolini, e-mail: ugolini@hp380.ist.unige.it
Received 30 Nov. 1998; revised 11 Mar. 1999; accepted 23 Apr. 1999.

Luukkonen defined citation analysis as the method that uses scientific criteria to measure the contribution of a published paper to the advancement of knowledge [9]. It is a widely debated approach, and many of the inherent drawbacks that spur disagreement have been pinpointed, namely, the selected journals included in databases providing the number of citations, the subjective motivations of citations or the language and publication type biases, multiple authorship merits and citing motivation [10, 11]. None the less, citation analysis remains a worthwhile criterion for evaluating the publication records of individual scientists, research units or national performance. The concept of citation analysis was originally proposed by Garfield [12] with the publication of the Science Citation Index (SCI), to which the Social Science Citation Index (SSCI) and the Arts and Humanities Citation Index (AHCI) have since been added. These indices define a citation as follows: when one document (a) mentions or refers to another document (b), the latter has been cited by the former as a source of information, as a support for a point of view, as authority for a statement of fact, etc. [13]. In practice, the term citation refers to the references listed at the end of a scientific article.

However, which indicators are appropriate, and how they should impact on the evaluation process or funding allocation decisions, are open questions. Despite the limitations mentioned above, bibliometrics may prove useful for attempts to gain insight into research product dissemination. With these considerations in mind, and taking into account the important medical and economic repercussions of cancer research, we performed an analysis of papers published in the journals listed by the Institute for Scientific Information (ISI) and correlated them with bibliometric parameters, i.e. impact factor (IF), and socioeconomic variables, i.e. the source country population and its gross domestic product (GDP). We also analysed the frequency of key words used in the oncological literature in order to identify the main research trends.

MATERIALS AND METHODS

Data of articles published in oncological journals listed by the ISI were selected in accordance with *Current Contents/Life Science* editorial and *Current Contents/Clinical Medicine* edition (1995–1997 actual years). Only bibliographic items with ISSN and nominal edition year 1995 were downloaded. Duplicate items were identified and deleted. All peer-reviewed papers, including editorials, reviews, technical notes and letters to the editor were included in this study. Journal supplements containing abstracts or meeting reports were excluded from the analysis.

For the purpose of this study, the definition of EU included the 15 official member states plus Norway, in view of its inclusion in the European Economic Area (EEA) and in all calculations concerning the EU issued by the Statistical Office of the European Communities (Eurostat). The country of the corresponding author was considered as the country of origin of the article. The papers originating from England, Scotland, Northern Ireland and Wales were grouped under the U.K. heading. For purposes of comparison, data on eight additional countries, each showing more than 40 entries during 1995, were also evaluated. On some occasions it was necessary to identify manually the country of origin of a given article after consulting other bibliographic databases. The country of origin of 3% of the articles, mainly editorials, remained unknown due to lack of specific data.

For the purpose of our study, key words were defined as comma-separated items of one or more words. All key words, both those reported by the authors and those attributed by the ISI, were identified and their frequency was calculated in two separate files using a special purpose program. Different key words with identical meaning were grouped and considered as a single key word. The same process was used for misspelled key words.

The resident population and GDP expressed in current billion US dollars for 1995 were retrieved for each country from Eurostat annual statistic reviews. Updated data are shown on the Internet site *www.cilnews.unige.it*.

Table 1. *Scientific productivity in surveyed countries (*1995*)*

	Oncological journals					Medical journals						
Country	No. papers	% (EU = 100)	Mean IF	Papers/GDP	Papers/million population	No. papers	% (EU = 100)	Mean IF	Papers/GDP	Papers/million population	% Oncology papers	IF ratio oncology/all
Austria	93	2.3	2.4	0.50	12	2166	1.9	2.6	11.7	270	4.3	0.9
Belgium	97	2.4	2.3	0.46	10	3367	2.9	2.6	15.9	337	2.9	0.9
Denmark	106	2.6	2.4	0.78	21	2893	2.5	2.4	21.2	562	3.7	1.0
Finland	91	2.2	2.6	0.96	18	2700	2.3	2.5	28.5	540	3.4	1.1
France	558	13.7	2.0	0.44	10	17 327	14.9	2.5	13.7	306	3.2	0.8
Germany	580	14.3	2.1	0.30	7	20 326	17.5	2.6	10.6	250	2.9	0.8
Greece	54	1.3	1.3	0.71	5	985	0.8	1.6	13.0	96	5.5	0.8
Ireland	26	0.6	2.0	0.59	7	821	0.7	2.1	18.5	233	3.2	0.9
Italy	761	18.7	2.2	0.65	13	11 243	9.7	2.4	9.6	197	6.8	0.9
Luxembourg	0	0	0	0.00	0	15	0	4.8	1.1	39	0	0
The Netherlands	410	10.1	2.9	1.31	27	7536	6.5	3.0	24.1	498	5.4	1.0
Norway	91	2.2	2.6	0.80	21	1666	1.4	2.2	14.7	388	5.5	1.2
Portugal	13	0.3	1.6	0.17	1	390	0.3	2.0	5.0	40	3.3	0.8
Spain	117	2.9	2.1	0.22	3	6535	5.6	2.2	12.4	169	1.8	0.9
Sweden	288	7.1	2.5	1.33	33	6322	5.4	2.5	29.3	727	4.6	1.0
U.K.	778	19.1	2.8	0.74	13	31 932	27.5	3.5	30.5	549	2.4	0.8
EU	4063	100	2.4	0.55	11	116 224	100	2.8	15.7	312	3.5	0.9
Australia	174	–	2.5	0.56	10	7206	–	2.6	23.3	409	2.4	1.0
Canada	344	–	2.9	5.36	13	12 316	–	3.0	191.8	451	2.8	1.0
India	100	–	1.4	0.41	<1	2685	–	1.6	11.0	3	3.7	0.9
Israel	109	–	2.1	1.58	22	3021	–	2.8	43.7	597	3.6	0.8
Japan	1127	–	2.4	0.29	9	23 136	–	2.4	5.9	186	4.9	1.0
Russia	41	–	1.5	0.12	<1	2079	–	1.1	6.0	14	2.0	1.4
Switzerland	119	–	2.5	0.47	17	4638	–	3.3	18.3	666	2.6	0.8
Taiwan	78	–	2.1	0.36	4	1712	–	1.9	7.8	82	4.6	1.1
U.S.A.	4523	–	3.3	0.70	17	118 565	–	4.0	18.4	455	3.8	0.8
World	11 117	–	2.7	0.46	2	309 684	–	3.1	12.9	62	3.6	0.9

RESULTS

Quantitative analysis

Overall, 11 117 papers were published in the oncological literature during 1995 (Table 1); of these, 4063 (36.5%) originated from the EU and 4523 (40.7%) from the US. All EU countries except Luxembourg were represented. The leading countries in Europe in terms of output were the U.K. (19.1%); Italy (18.7%); Germany (14.3%); France (13.7%); and The Netherlands (10.1%).

For purposes of comparison, during the same year a total of 309 684 papers were published in the whole of measurable world medical literature; 116 224 (37.5%) were by EU authors, and 118 565 (38.3%) by US authors. Oncological papers accounted for 3.6% of the total number of medical papers (Table 1). In the EU, the percentage of oncological papers was 3.5% and in US was somewhat higher (3.8%). The ratio of oncological papers to all medical literature was highest in Italy (6.8); Greece (5.5); Norway (5.5); The Netherlands (5.4); Sweden (4.6); Austria (4.3); Denmark (3.7); Japan (4.9); Taiwan (4.6); and India (3.7).

Qualitative analysis

The mean IF of papers from the EU in oncological journals was nearly 2.4 in comparison with 3.3 for the US (Table 1). The world IF for oncological papers was 2.7. Among EU nations, The Netherlands ranked first with a mean IF of 2.9, followed by the U.K. (2.8), Finland and Norway (2.6); Sweden (2.5); and Austria and Denmark (2.4). The mean IF in other non-EU countries ranged from 1.4 for India to 2.9 for Canada. The mean IF of all medical literature produced in the EU was higher (2.8) than that of oncological papers. Luxembourg (4.8), the U.K. (3.5) and The Netherlands (3.0) had particularly high total medical literature IF. The ratio between the IFs of oncological journals and all medical literature was calculated. A value exceeding one, indicating that in a given country oncological papers reached a higher IF than the rest of medical literature, was shown by Russia (1.4); Norway (1.2); Finland (1.1); and Taiwan (1.1).

GDP and population size comparison

The ratio between the number of published oncological papers and GDP showed a mean value of 0.55 for the EU, a finding which compared unfavourably with that of 0.70 calculated for the US (Table 1). In the EU, Sweden ranked first (1.33), followed by The Netherlands (1.31); Finland (0.96); Norway (0.80); and Denmark (0.78). When all medical scientific literature was considered, the US scored better than the EU (18.4 versus 15.7). The U.K. (30.5), Sweden (29.3), Finland (28.5) and The Netherlands (24.1) showed the highest values in the EU. The highest value belonged to Canada, with 191.8 papers/billion US dollars of GDP.

The ratio between the number of published oncological papers and country's population in millions was 11 for the EU and 17 for the US (Table 1). In Europe, small countries generally performed better than larger ones. Sweden ranked first with the best world score (33), followed by The Netherlands (27), Norway and Denmark (21) and Finland (18). Outside the EU and U.S.A., high scores were seen for Israel (22). The analysis of the world medical output by a country's population yielded values of 312 in the EU and of 455 in the US. The highest EU scores were those of Sweden (727), Denmark (562), the U.K. (549) and Finland (540). Switzerland (666) and Israel (597) ranked high among non-EU countries.

Research topics

In oncological journals, the key words attributed by authors comprised as many as 18 209 different terms. Of these, only 9784 were cited more than twice, and 5100 were cited more than 10 times. 16 064 key words attributed by ISI appeared in the oncological literature. Of these, 10 076 were cited more than twice, and 5465 more than 10 times. Misspelled or non-standardised key words were found frequently. Indeed, this lack of standardisation of key words greatly hampered our analysis; in fact, the use of generally accepted synonymous terms (e.g. cancer versus carcinoma, neoplasm versus neoplasia, apoptosis versus cell death versus programmed cell death, be these expressed either as plural or singular terms), of spelling variations (e.g. tumor versus tumour), of generic rather than specific terminology (e.g. oncogene(s) versus c-myc) and of abbreviations (particularly for drug names, such as 5-fluorouracil, which appeared in six different forms) leads to a great deal of disparity in the attribution of key words, and consequently, difficulty in the interpretation of an author's exact intended meaning. Nevertheless, the analysis of key words allowed us to assemble and arrange those with similar meaning in order to produce a list of the most often cited terms: the top ten key words related to disease types, those related to drug types and those related to research topics and related techniques are listed in Table 2.

Table 2. Analysis of keywords for oncology publications

	Number of occurrences
Diseases	
Breast cancer	361
Prostate cancer	223
Colorectal cancer	200
Lung cancer	135
Head and neck cancer	125
Uterine cancer	120
Gastric cancer	96
Ovarian cancer	88
Glioma	74
Melanoma	67
Drugs	
Cisplatin	100
Interleukins	65
5-Fluorouracil	60
Taxol and analogues	46
Monoclonal antibodies	34
Tamoxifen	32
Interferons	30
Doxorubicin/epirubicin	42
Retinoic acids	24
Cyclophosphamide/ifosfamide	20
Research topics and related techniques	
p53	249
Metastasis	111
Growth factor	106
Immunohistochemistry	99
Proliferation	73
Apoptosis	65
ras	64
Tumour suppressor gene	64
PCR	61
Oncogene	46

DISCUSSION

Our data show that the geographic breakdown of the oncological literature output has scarcely changed with respect to the values reported in 1993 by Parodi and colleagues [8]. US and U.K. researchers still produce both the most medical and oncological literature. It should be pointed out that bibliometric analyses are biased towards English language journals, and authors of some nations (e.g. France and Germany) with a strong tradition of publishing in their native languages and less prone to submitting papers to internationally peer-reviewed English-language journals may be penalised in comparative studies drawing on databases that include few non-English-language publications. For example, the database *EMBASE* includes more non-English-language journals than does *Medline* [14]. None the less, English has become the common language of the international scientific community, and future bibliometric studies are likely to be performed only on English-language journals.

The oncological literature output in the EU was somewhat lower than that of the US in terms of number of papers, despite a large difference in research fundings available across the Atlantic, and is in line with overall medical literature production. This finding confirms that the EU, which is increasingly becoming an integrated geographic area, plays a leading role in oncological research. However, the mean IF of oncological papers was higher in the US than in the EU.

In 1995, authors from every EU country, except Luxembourg, published papers in the oncological journals considered here. The U.K., Italy, Germany and France were the top four ranking countries for total number of published papers. This ranking changed considerably when other variables, such as mean impact factor, number of papers per inhabitants, or number of papers for GDP, were assessed. The Netherlands, U.K., Finland and Norway excelled for mean impact factor. Only the U.K. held a good position for IF, confirming that this country is still a leader in scientific production in Europe. Sweden and The Netherlands had the highest scores for the ratio between scientific publications, number of inhabitants and GDP. Our results are consistent with the view that smaller countries usually perform better than larger ones in terms of cancer research literature output. These data also confirm findings published in *Science* in 1992 [15] regarding the overall scientific output between 1981 and 1990 in Europe: the U.K., Germany, France and Italy ranked first to fourth for number of papers published, whilst Switzerland, Sweden, Denmark and The Netherlands were the countries with the highest number of mean citations per paper. Other works have documented similar trends. Benzer and colleagues compared the overall number of medical publications in 1990 [16], and found that Israel, Sweden and Switzerland had the highest scientific output per country population. In another study, where the number of papers produced by a given country was related to the number of physicians [17], the U.K. was first with 0.37 publications per physician, followed by Israel, Switzerland, Denmark, Finland, Sweden and The Netherlands. Finally, Hausen and colleagues [18] measured scientific production according to the gross domestic expenditure for research in 1989, and found that New Zealand, Denmark, Spain, Canada and The Netherlands scored the highest.

The reason why smaller countries have a higher scientific output than larger ones is not known, but some explanations, for instance, better utilisation of resources, a higher percentage of the GDP allocated to research, a clustering of specific oncological diseases, have been advanced. Our analysis of key words revealed a high degree of disparity of terms used in oncological journals, with only 28% of key words being cited more than 10 times, and 53.7% cited more than twice. This fact was due mainly to the use of different synonyms, and only marginally to misspelling. This problem affects many medical disciplines, and in some cases to a greater extent than oncology: in fact, our data show that the percentages of key words cited more than 10 times in rheumatological, dermatological or geriatric journals in 1995 were 2.1, 3.6 and 2.8%, respectively [19]. Given findings such as these, we believe that editors should adopt measures that encourage the standardisation of key words, in order to facilitate the retrieval of bibliographical information from computerised sources. A standardised key word system would ideally include items such as field of research, disease and its localisation, methods implemented, and relevant drugs or biologicals used. Our data show that the most commonly used key words for diseases very closely reflects the incidence of the numerous types of neoplastic diseases [20], with breast cancer ranking the top. Among drugs cited, interest focuses on the chemotherapeutic drugs most widely used in current clinical settings, but substances such as cytokines, monoclonal antibodies, paclitaxel or retinoic acids also emerge as widely studied subjects. Finally, the prevailing trend of experimental studies is the widespread application of the molecular biology to experimental oncology studies in terms of both phenomena studied and techniques used.

Admittedly, our study took into account only those journals classified as oncological, and did not endeavour to tackle the complicated task of identifying the wealth of cancer research studies appearing in other scientific publications. This limits data evaluation to only a part of all oncological productivity since articles on basic cancer research may be published in basic discipline journals (e.g. biochemistry, immunology) and articles on clinical cancer studies may be published in categories covering general subjects (e.g. medicine, pharmacology, etc.) or affected systems or organs (e.g. respiratory systems, digestive systems). Ideally, an exhaustive survey of cancer research output would seek to combine data on both key words identifying the subject of analysis and authors' affiliation, tasks respectively entailing the need to search and assess different databases with different classification indexing schemes and to handle the inaccurate reporting of authors' addresses. Our intention, however, was not to provide an exhaustive survey of cancer research output, but rather to offer a broad review of the data and to gather impressions of publication trends, in comparison with other studies on the subject.

The two main problems that our analysis has evinced are the lack of standardisation of key words and the inaccurate reporting of corresponding authors' addresses. We are currently developing a method to identify the scientific background of authors that may contribute to overcoming these matters, whereby key words, names of the authors and relevant affiliations are matched. Measuring the impact of research is a difficult process that is currently the focus of significant debate throughout the world. This evaluation exercise should also take into consideration the effect on research targets (technology, systems, education, social structure, etc.) and all expressions of knowledge (patents and trained students) in addition to published papers. Whilst some impacts are tangible, many are intangible and difficult to identify and quantify. Assessing the impact of a scientific paper by summing the number of times it is referenced by other authors is obviously a method that is neither flawless nor invulnerable and provides an incomplete picture of the research product. Despite its limitations, however, many studies have shown that this approach provides useful information for the evaluation of productivity, and may well constitute a good general guideline in an era of cost-effectiveness and quality control. Though there is little evidence that the results from such studies are actually used in practice by research evaluators and funding agencies, there is no doubt that a country could benefit from their application, in order to determine its position with respect to competitors and, in turn, to exploit opportunities arising in all scientific fields.

1. Davis RA, Cunningham PS. Creative thought in neurosurgical research: the value of citation analysis. *Neurosurgery* 1990, **26**, 345–353.
2. Reed KL. Citation analysis of faculty publication: beyond Science Citation Index and Social Science Citation Index. *Bull Med Libr Assoc* 1995, **83**, 503–508.
3. Schena FP, Zoccali C. Bibliometric analysis of the scientific productivity of the Italian Society of Nephrology for a period of five consecutive years (1990–1994). *Nephrol Dial Transplant* 1996, **11**, 2359–2360.
4. Ell C, Gunreben E, Kettner S, Lehrl S. Research activity and research quality of German gastroenterologists reflected by citation indices. *Z Gastroenterol* 1997, **35**, 23–28.
5. Hansen HB, Brinck K, Henriksen JK. Scientific publications from departments of clinical physiology and nuclear medicine in Denmark. A bibliometric analysis of "impact" in the years 1989–1994. *Clin Physiol* 1996, **16**, 507–519.
6. Adams J, Griliches Z. Measuring science: an exploration. *Proc Natl Acad Sci USA* 1996, **93**, 12664–12670.
7. Bencetic Klaic Z, Klaic B. Scientometric analysis of anthropology in the Republic of Croatia for the period 1980–1996. *Coll Antropol* 1997, **21**, 301–318.
8. Parodi S, Parodi A, Lombardo C, Santi L. Cancer research in the European Community and other non-EC Countries. *Tumori* 1993, **79**, 9–15.
9. Luukkonen T. Bibliometrics and evaluation of research performance. *Ann Med* 1990, **22**, 145–150.
10. MacRoberts MH, MacRoberts BR. Problems of citations analysis. *Scientometrics* 1996, **36**, 435–444.
11. Schoonbaert D, Roclants G. Citation analysis for measuring the value of scientific publications: quality assessment tool or comedy of errors? *Trop Med Int Health* 1996, **1**, 739–752.
12. Garfield E. Citation indexes for science: a new dimension in documentation through association of ideas. *Science* 1955, **122**, 108–111.
13. Institute for Scientific Information. *SCI: Science Citation Index 1996*. Philadelphia, U.S.A, The Institute for Scientific Information, 1996.
14. Staab MA. Geography of medical publications. *Lancet* 1993, **341**, 634.
15. Anderson A. Science in Europe. *Science* 1992, **256**, 472.
16. Benzer A, Pomaroli A, Hauffe H, Schmutzard E. Geographical analysis of medical publications. *Lancet* 1993, **341**, 247.
17. Eldor J. Geography of medical publications. *Lancet* 1993, **341**, 634.
18. Hausen A, Reibnegger G, Wachter H. Geography of medical publications. *Lancet* 1993, **341**, 634.
19. Mela GS, Cimmino MA. An overview of rheumatological research in the European Union. *Ann Rheum Dis* 1999, 1998, **57**, 643–647
20. In Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J, eds. Cancer incidence in five continents. *Vol VII*. Lyon, IARC Scientific Publications no. 143, 1997.

Acknowledgements—We are grateful to Mr Thomas Wiley for reviewing the English format of the manuscript.