Industrial Health 2010, **48**, 197–203 **Original Article**

Density-Equalizing Mapping and Scientometric Benchmarking in INDUSTRIAL HEALTH

Cristian SCUTARU[1, 2]*, David QUARCOO[1], Masaya TAKEMURA[1], Tobias WELTE[2], Tanja C. FISCHER[3] and Beatrix GRONEBERG-KLOFT[2, 4]

[1]Department of Information Science, Institute of Occupational Medicine, Charité-Universitätsmedizin Berlin, Free University Berlin and Humboldt-University Berlin, Thielallee 73, 14195, Berlin 14195, Germany
[2]Department of Respiratory Medicine, Hannover Medical School, Hannover, Germany
[3]Allergy-Centre-Charité, Department of Dermatology and Allergy, Charité-Universitätsmedizin Berlin, Free University Berlin and Humboldt-University Berlin, Berlin, Germany
[4]Otto-Heubner-Centre, Charité-Universitätsmedizin Berlin, Free University Berlin and Humboldt-University Berlin, Berlin, Germany

Received February 3, 2009 and accepted September 2, 2009

Abstract: Bibliometric techniques have been introduced to the field of industrial health in the past two decades. Since then, several studies have assessed progression of science in this area using quantitative measures and qualitative measures such as impact factor or H-indices. Since novel procedures such as density-equalizing mapping have not been used so far, the present study combined classical bibliometric tools with novel scientometric and visualizing techniques. All "INDUSTRIAL HEALTH" entries listed in the ISI database since 1987 were screened and analyzed. Using bibliometric approaches, a continuous increase in qualitative markers such as collaboration numbers or citations were found while quantity markers such as author numbers or publication numbers remained relatively constant. The combination with density equalizing mapping revealed a distinct global pattern of research productivity and citation activity with Japanese institutions at the leading position. Radar chart techniques were used to visualize bi- and multilateral research cooperations and institutional cooperations. In summary, the present study supplies a first scientometric-bibliometric approach that visualizes research activity in "INDUSTRIAL HEALTH" over the past decades.

Key words: Scientometry, Density equalizing, INDUSTRIAL HEALTH, Occupational health, Radar chart

Introduction

Smith *et al*. and Sawada *et al*. recently addressed the issue of bibliometric research in occupational health[1, 2]. It was elegantly demonstrated that bibliometrics — which is defined as the use of mathematical techniques to investigate publishing and communication patterns in the distribution of information — has been an established approach in occupational and industrial health for about two decades. By this time, McCunney and Harzbecker published a study on citation patterns[3]. This study was followed by other reports by Takahashi *et al*.[4], Sizaret and Kaufmann[5], and Gehanno *et al*.[6] leading to new proposal for improved measures of quality indices by Takahashi *et al*.[7], Garfield[8] and Gehanno and Thirion[9].

The focus of these studies was either to supply bibliometric data or to analyse and improve bibliometric procedures in the field of occupational health. While the gain of knowledge concerning these studies was recently summarized by Smith *et al*.[1] and there are also reports on publishing trends and citation indexing available[10, 11], novel scientometric techniques in combination with visualizing techniques have not been implemented so far. Therefore, the present study was designed to visualize research activity using INDUSTRIAL HEALTH articles and density equalizing procedures in combination with classical bibliometric techniques in accordance to the

*To whom correspondence should be addressed.
E-mail: cristian.scutaru@charite.de

NewQIS protocols[12, 13].

Subjects and Methods

Data source and time span

Data was retrieved from the database Web of Science database from Thomson Reuters[14, 15] as previuously described[16]. As restriction concerning the publication date, the period between 1987 (inclusion to Web of Science database) and 2008 (last completed year) was used.

Search strategies

All published items in the journals "INDUSTRIAL HEALTH" were selected using the web interface. No additional filters were used.

"Quality" parameter analysis

For countries with at least 30 published items in "INDUSTRIAL HEALTH", the average citation per published article was calculated.

Secondly, the theory of the H-index was extrapolated to the articles originating from a specific country and a similar country specific modified H-index was calculated in order to assess the "quality" of articles from a specific country. The original H-index for authors defined by Hirsch, is a quality index (h), where h is the maximum number from the total number of articles written by a given author where each one of these h articles have been at least h times cited[17].

Density-equalizing mapping

Density-equalizing mapping was used as described in previous studies[18]. In brief, the territories were re-sized according to a particular variable, i.e. the number of published items of Industrial Health, the total citations or a modified country H-index, respectively. For the re-sizing procedure the area of each country was scaled in proportion to the variable. The calculations of the procedure are based on Gastner and Newman's algorithm[19].

Analysis of multilateral country and institution cooperations

A bilateral cooperation between 2 countries was defined when at least one author originates from one country and at least one other author from a second country. A matrix with all participant countries was computed with a special software and filled with the appropriate values for the cooperation for each pair of countries. A second module of this software was developed to interpret the matrix and transform the figures into vectors. The thickness of the vector quantifies the number of cooperation articles between the two countries. A threshold of at least two cooperations was set in order to improve the readability.

In parallel, a cooperation network of the publishing institutes was computed and graphically visualised using a radar chart. For this visualisation a threshold of at least three cooperations was set in order to improve the readability.

Results

Quantitative parameters: Total number of published items and average authorship in INDUSTRIAL HEALTH

The number of published items was used as an index of quantity of research productivity. In total, a number of 1,095 INDUSTRIAL HEALTH articles were included in the Web of Science database. The first articles were published in 1987. The year 2007 holds the largest number of published items (113), followed by 2006 (100) and 2008 (85) (Fig. 1).

Japan is the country with the highest output with a total of 691 articles. The United States are in second place with 68, followed by South Korea and India with each 63, Italy (30), China and Taiwan (29), United Kingdom (24), Germany (23) and Thailand (18).

When analyzing for average number of authors per publication over the time, very steady values were recorded ranging between 3.31 authors per publication in 1987 to a maximum of 4.68 authors per publication in 2007 (Fig. 1).

Density equalizing mapping of research output in INDUSTRIAL HEALTH

Density-equalizing mapping was used according to a recently published method to illustrate the research output (number of published articles in INDUSTRIAL HEALTH) by territorial resizing. As evident from the total article analysis, it is obvious that the cartogram is dominated by Japan with an output of 691 articles. This is nearly 10 times the output of the countries ranked next (United States and South Korea and India (Fig. 1)).

Country research network analysis in INDUSTRIAL HEALTH

There is an overall increase in international research cooperations present. In 1987, the number of international cooperations was below 5 (Fig. 2A). The year 2008 holds the largest number of cooperation articles (23), followed by 2006 (17) and 2007 (14) (Fig. 2A).

To visualize research networking for INDUSTRIAL HEALTH articles, the radar chart technique was used and it was found that with a number of 17 cooperation articles Japan and China are the leading cooperating countries. This is followed by the cooperations of Japan and the United States (16) and Japan and South Korea (15) (Fig. 2 B).

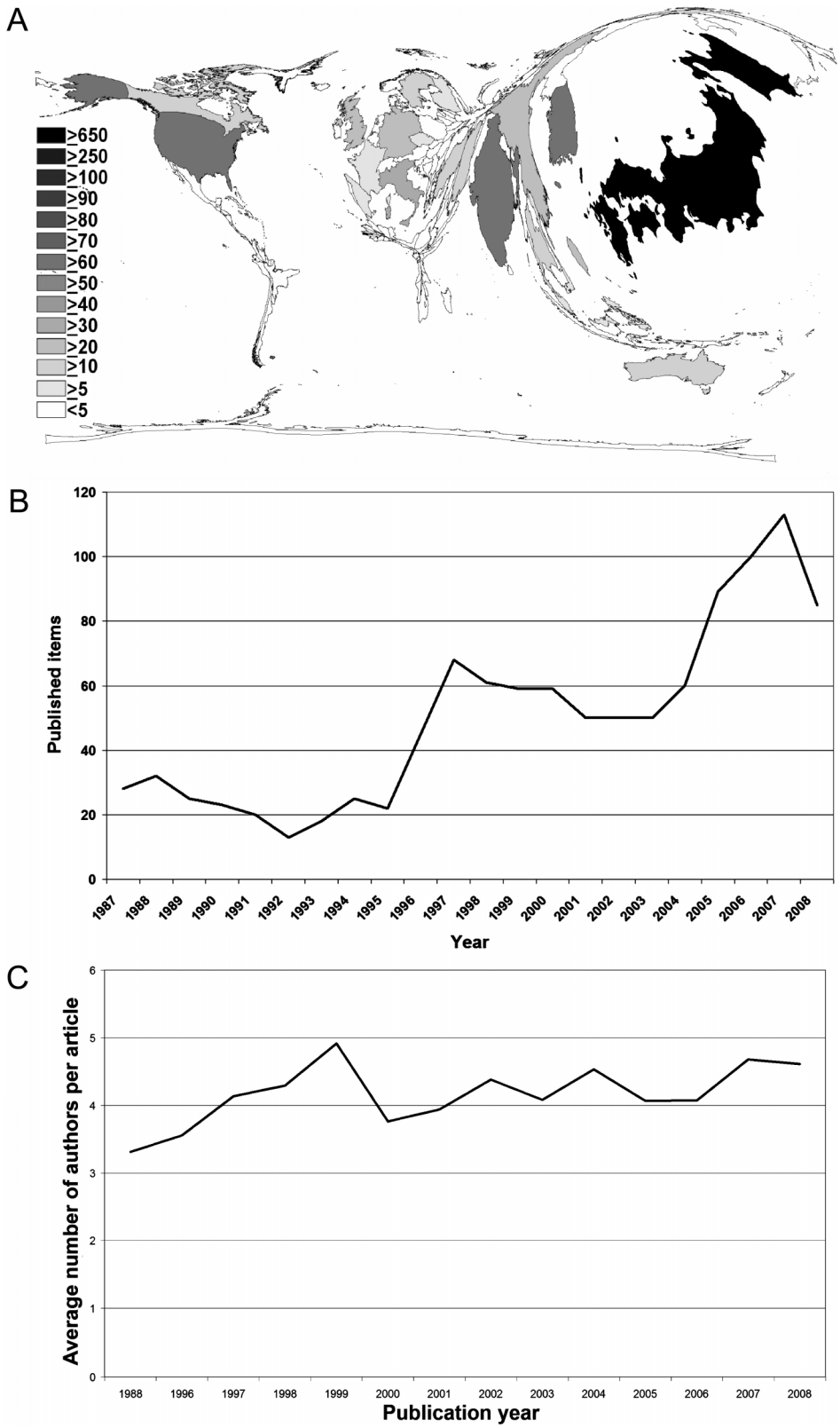

Fig. 1. A: Density-equalizing map illustrating the number of contributions for each country in INDUSTRIAL HEALTH for the period 1987–2008.

The area of each country was scaled in proportion to its total number of publications. Colors encode the number of contributions per country. B: Evolution of article numbers in the period 1987–2008. 2008 entries not completely finished. C: Average number of authors per article over the time.

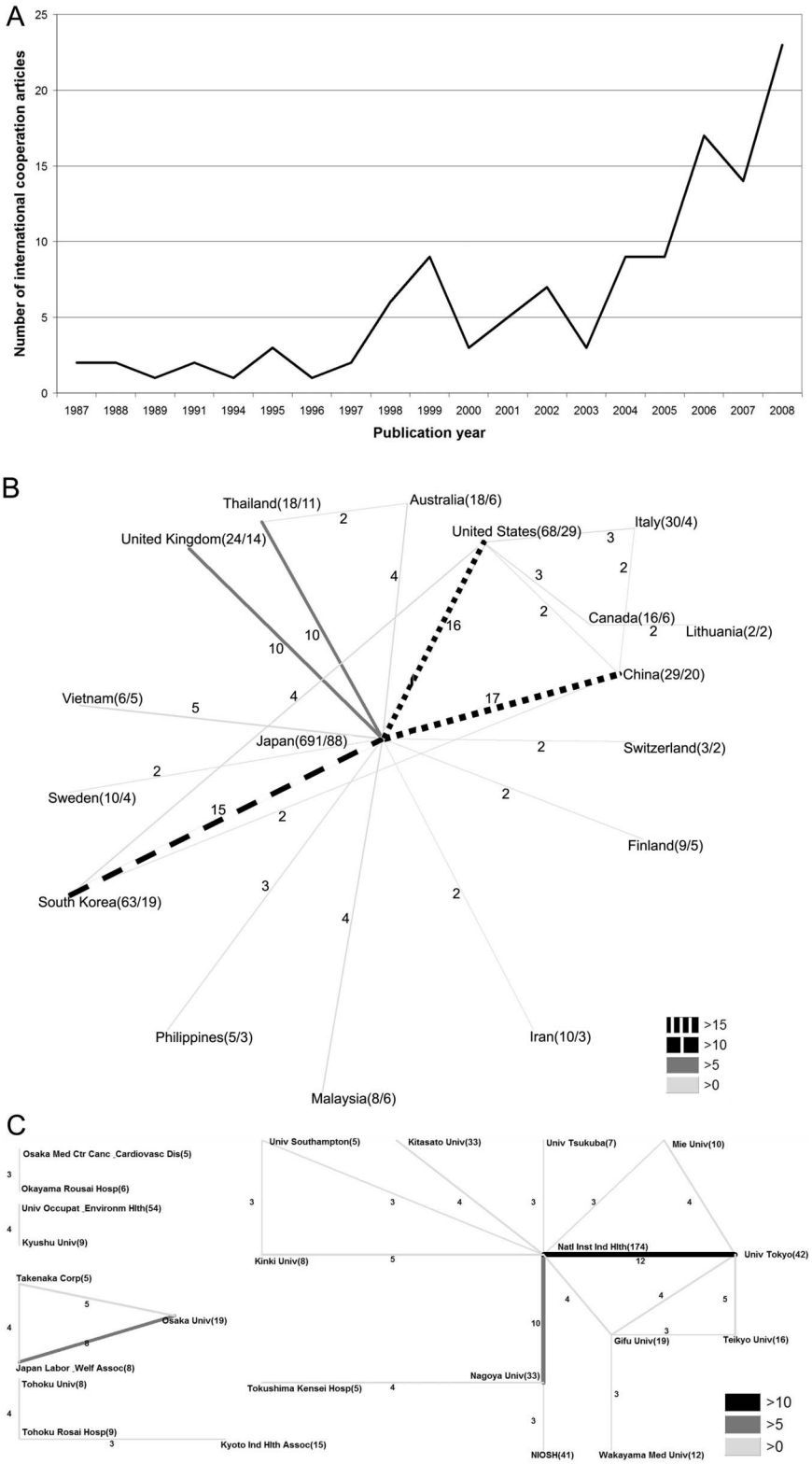

Fig. 2. Network analysis.

A: Country network analysis. Evolution of the number of international cooperations between 1987 and 2008. B: Radar chart visualizing bilateral networking between countries for the overall number of cooperations. Greyscale and size of bars encode the number of bilateral cooperations. C: Institutional network analysis using the radar chart technique. Greyscale and size of bars encode the number of bilateral cooperations.

Institutional research network analysis in INDUSTRIAL HEALTH

In order to identify and visualize leading institutional networks the radar chart technique was applied (Fig. 2C). The highest cooperation value was present for cooperations between the institutions "Natl Inst Ind Hlth" and "Univ Tokyo", with a number of 12 articles. In second place cooperations between "Natl Inst Ind Hlth" and "Nagoya Univ" were found (10). The third highest cooperation number was present for "Osaka Univ" and "Japan Labor & Welf Assoc" with 8 published items (Fig. 2C).

"Quality" parameter analysis: citations and modified country H-index

In order to assess qualitative parameters, citations and the modified country H index were calculated and visualized by density equalizing mapping. In total, INDUSTRIAL HEALTH articles originating from Japan gathered the highest number of citations (3,122). South Korea was placed in second with 285 citations, followed by the United States with 277 citations. This information was visualised by means of density equalising mapping in Fig. 3A.

Only 5 countries had at least 30 published items. When a ranking for these countries concerning the average citation per item was established, relatively similar values were found for Italy (5.93), South Korea (4.52), Japan (4.51), the United States (4.07) and India (4.06).

As a further "quality" parameter, the modified country H-index was used. In this analysis, Japan reached the highest value (19) followed by the United States (10), South Korea (10) and India (8) as shown in Fig. 3B.

Discussion

As elegantly described in recent articles by Smith and Gehanno *et al*., there has been an increasing level of debate regarding the overall usefulness and relevance of impact factors for occupational health[20, 21]. One comprehensive, general review article on this topic has been published by Smith in 2007[22]. The present study is the first analysis to assess the scientific progress in INDUSTRIAL HEALTH using a combination of novel visualizing tools such as density equalizing mapping and classic bibliometric tools such as publication and citation analysis.

An increasing number of networks was found when the data of the different years since 1987 were compared. This needs to be interpreted in the context of occupational health research and funding: Occupational and environmental diseases are estimated to exert a major burden of disease, with most of the costs due to both diagnosis and therapy and reduced productivity at work[23–25]. Therefore, numerous national and international research networks were founded by governmental and non-governmental institutions[26]. These research networks are at least partly responsible for the increasing number of multilateral cooperations which have been found for the INDUSTRIAL HEALTH publications in the present study.

For the present study, it is important to realise that the analysis of INDUSTRIAL HEALTH should not be regarded as completely representative for global occupational research activity. In this respect, a bias is represented by the host country of the Journal, Japan. Quantitative data analysis of productivity parameters shows that research groups from Japan maintain a leadership position for INDUSTRIAL HEALTH articles since 1987. This is a normal finding which is also presented for other journals and other countries. This was also evident after visualization by the density equalizing mapping technique using the Gastner and Newman's algorithm[19]. This technique also illustrates that there is a large number of non-Japanese institutions that contribute to INDUSTRIAL HEALTH which can therefore be clearly defined as an international journal. In this respect, the international orientation of the journal is not only represented by the journal's language but also by the multitude of nations contributing with the United States ranked second followed by South Korea, India, Italy, China and Taiwan, United Kingdom, Germany and Thailand.

Whereas the number of published items was currently considered as an index of quantity of research productivity, citation analysis may be used as an indicator for research quality. In general, quality indicators need to be regarded extremely critically and therefore, not to be over interpreted as indicated by numerous previous articles[7–9].

We assessed citations and the modified country H index and visualized the scores by density equalizing mapping. Parallel to the quantity parameter, articles originating from Japanese institutions gathered the highest total number of citations. In the average citation per item analysis that is usually dominated by countries with only a few contributions, relatively similar values were found for Italy, South Korea, Japan, the United States and India. This indicates a uniform distribution of the research quality. As a further parameter, the modified country H-index was used and Japan reached the highest value indicating the high quality of research in terms of citability.

Conclusion

The present study represents the first visualisation of INDUSTRIAL HEALTH articles using density equalizing calculations and radar chart techniques.

The data demonstrates an international journal with a strong increase in research productivity and networking

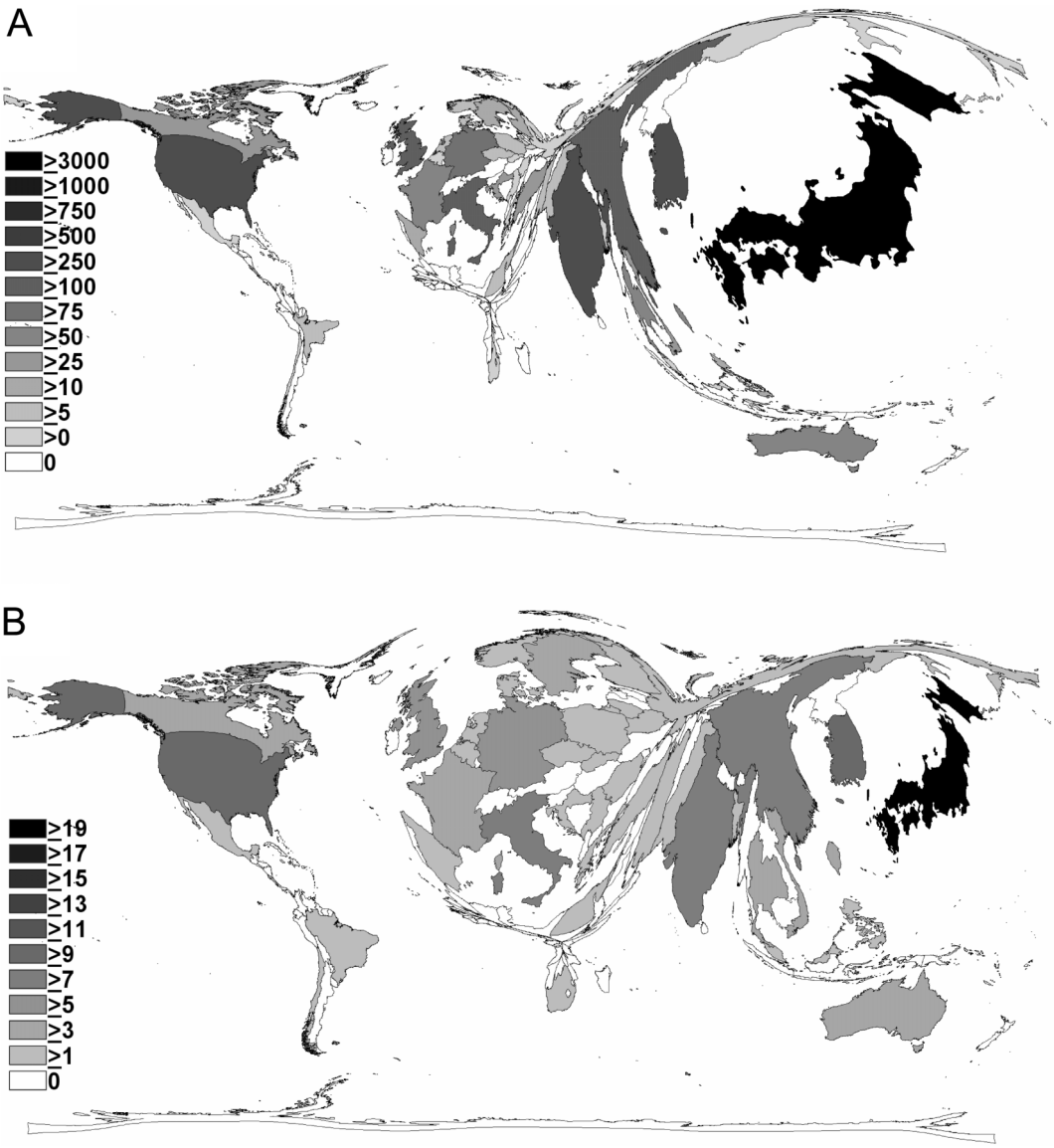

Fig. 3. "Quality" assessment.
A: Density-equalizing map illustrating the number of total citations for each contributing country in INDUSTRIAL HEALTH for the period 1987–2008. The area of each country was scaled in proportion to its total number of publications. Colors encode the total number of citations per country. B: Density-equalizing map illustrating the modified country H-index for each contributing country in INDUSTRIAL HEALTH for the period 1987–2008. The area of each country was scaled in proportion to its modified H-index. Colors encode the modified H-index per country.

over the past years. The techniques established here can be of use for future bibliometric studies in occupational medicine.

References

1) Smith DR, Gehanno JF, Takahashi K (2008) Bibliometric research in occupational health. Ind Health **46**, 519–22.
2) Sawada S, Smith DR, Araki S (2007) The impact factor and INDUSTRIAL HEALTH. Ind Health **45**, 501–2.
3) McCunney RJ, Harzbecker J (1992) The influence of occupational medicine on general medicine: a look at the journals. J Occup Med **34**, 279–86.
4) Takahashi K, Hoshuyama T, Ikegami K, Itoh T, Higashi T, Okubo T (1996) A bibliometric study of the trend in articles related to epidemiology published in occupational health journals. Occup Environ Med **53**, 433–8.
5) Sizaret A, Kauffmann F (1997) A bibliometric study of the trend in articles related epidemiology published in occupational health journals. Occup Environ Med **54**, 357–8.
6) Gehanno JF, Paris C, Thirion B, Caillard JF (1998)

Assessment of bibliographic databases performance in information retrieval for occupational and environmental toxicology. Occup Environ Med **55**, 562–6.
7) Takahashi K, Aw TC, Koh D (1999) An alternative to journal-based impact factors. Occup Med (Lond) **49**, 57–9.
8) Garfield E (1999) Refining the computation of topic based impact factors—some suggestions. Occup Med (Lond) **49**, 571–2.
9) Gehanno JF, Thirion B (2000) How to select publications on occupational health: the usefulness of Medline and the impact factor. Occup Environ Med **57**, 706–9.
10) Smith DR, Sawada S, Araki S (2007) Twenty years of publishing trends and citation indexing at INDUSTRIAL HEALTH, 1987–2006. Ind Health **45**, 717–20.
11) Smith DR, Leggat PA (2009) Estimation of some missing bibliometric indicators at INDUSTRIAL HEALTH. Ind Health **47**, 202–3.
12) Groneberg-Kloft B, Fischer TC, Quarcoo D, Scutaru C (2009) New quality and quantity indices in science (NewQIS): the study protocol of an international project. J Occup Med Toxicol **4**, 16.
13) Groneberg-Kloft B, Quarcoo D, Scutaru C (2009) Quality and quantity indices in science: use of visualization tools. EMBO Rep **10**, 800–3.
14) Sevinc A (2005) Multilingual approach to "Web of Science". J Natl Med Assoc **97**, 116–7.
15) Sevinc A (2004) Web of science: a unique method of cited reference searching. J Natl Med Assoc **96**, 980–3.
16) Groneberg-Kloft B, Dinh QT, Scutaru C, Welte T, Fischer A, Chung KF, Quarcoo D (2009) Cough as a symptom and a disease entity: scientometric analysis and density-equalizing calculations. J Investig Allergol Clin Immunol **19**, 266–75.
17) Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A **102**, 16569–72.
18) Groneberg-Kloft B, Scutaru C, Kreiter C, Kolzow S, Fischer A, Quarcoo D (2008) Institutional operating figures in basic and applied sciences: scientometric analysis of quantitative output benchmarking. Health Res Policy Syst **6**, 6.
19) Gastner MT, Newman ME (2004) From The Cover: Diffusion-based method for producing density-equalizing maps. Proc Natl Acad Sci U S A **101**, 7499–504.
20) Smith DR (2008) Citation analysis and impact factor trends of 5 core journals in occupational medicine, 1985–2006. Arch Environ Occup Health **63**, 114–22.
21) Gehanno JF, Takahashi K, Darmoni S, Weber J (2007) Citation classics in occupational medicine journals. Scand J Work Environ Health **33**, 245–51.
22) Smith DR (2007) Historical development of the journal impact factor and its relevance for occupational health. Ind Health **45**, 730–42.
23) Fingerhut M, Nelson DI, Driscoll T, Concha-Barrientos M, Steenland K, Punnett L, Pruss-Ustun A, Leigh J, Corvalan C, Eijkemans G, Takala J (2006) The contribution of occupational risks to the global burden of disease: summary and next steps. Med Lav **97**, 313–21.
24) Driscoll T, Nelson DI, Steenland K, Leigh J, Concha-Barrientos M, Fingerhut M, Pruss-Ustun A (2005) The global burden of disease due to occupational carcinogens. Am J Ind Med **48**, 419–31.
25) Friedman C, McKenna MT, Ahmed F, Krebs JG, Michaud C, Popova Y, Bender J, Schenk TW (2004) Assessing the burden of disease among an employed population: implications for employer-sponsored prevention programs. J Occup Environ Med **46**, 3–9.
26) Cullen MR, Upton A, Buffler P, Robins T, Schenker M, Fine L, Wiencek R, Widess E, Chiazze L Jr (1994) The private funding of public research. New directions in the administration of occupational and environmental health research. J Occup Med **36**, 1348–54.