

The Non-Source Items of the Journal Citation Reports

Johannes Stegmann* and Guenter Grohmann**

*johannes.stegmann@charite.de

Charité - University Medicine Berlin, Campus Benjamin Franklin, Medical Library, D-12203 Berlin (Germany)

**guenter.grohmann@charite.de

Charité - University Medicine Berlin, Campus Benjamin Franklin, Institute of Medical Informatics, Biometry and Epidemiology, D-12203 Berlin (Germany)

Introduction

The two editions of the Journal Citation Reports (JCR), published annually by the Institute for Scientific Information (ISI), list all the journals which are established sources of ISI's article databases. The JCR science edition 2003 includes 5907 journal titles, the respective social sciences edition lists 1714 titles, classified to roughly 200 different scientific subfields or categories (Institute for Scientific Information, 2003). A main feature of the JCR is the presentation of journal impact factors which relate the number of cites a journal receives to the number of its published papers and serve as a measure of a journal's scientific "weight". By use of the category scheme of the JCR it is possible to calculate subfield-specific impact factors which are the basis of any serious comparison between distinct institutions and/or subfields.

It is well known that the total number of scientifically relevant journals is far beyond the number of titles included in the JCR (Allee, 1987; Garfield, 1996; Schoonbaert & Roelants, 1996; Stegmann, 1999). Several attempts were undertaken to calculate impact factors for at least some of those journals which are cited by ISI source journals but are not included in the JCR (Spaventi et al, 1979; Sen, Karanjai & Munshi, 1989; Christensen, Ingwersen & Wormell, 1996; Stegmann, 1997, 1999). Recently, Butler & Visser (2004) reported preliminary results regarding the extent to which extent that part of the scientific output of Australian universities which is not indexed in ISI databases is cited by ISI-source titles, and Visser & Moed (2004) proposed a general framework how to identify and deal with cited ISI-non source items in studies concerning disciplines with lower ISI coverage. For this communication the current JCR editions (2003) were inspected to figure out the extent to which the JCR source journals and the various field-specific categories constituted by them cite references outside the ISI world.

Methods

The *Citing Journal Listings* of the JCR list the abbreviated source titles of all material cited by the respective citing (JCR-included) journal. Thus, it is

possible to detect by proper matching procedures the cited sources which are included in the JCR as citing journals as well as those which are not included. We call the former "source titles" and the latter "non-source items". From the *Citing Journal Listings* the number of cites given in 2003 to source titles and non-source items published in 2001/2002 were extracted and assigned to the JCR categories to which the citing journals belong. Only cited items represented by an (abbreviated) title string were taken into account. The cites given to *ALL OTHERS* were omitted.

Results

The science edition of the JCR 2003 counts about 3.37 million cites. 92.4% of them were given to JCR source titles, 7.6% to non-source items. The data for the respective JCR social sciences edition are: 65.3% of nearly two hundred thousand cites given to source titles, 34.7% given to non-source items. Consequently, most of the subfields of the social sciences edition of the JCR cite non-source items to a large extent. Only one field gives less than 10% of its cites to non-source items (category PSYCHOLOGY, BIOLOGICAL, 6.9%). The categories PSYCHIATRY and NURSING, e.g., cite non-source items to an extent of 10.3% and 32.6%, respectively. The category AREA STUDIES shows the highest value (81.2%) for cites given to titles outside ISI sources.

The science edition of the JCR 2003 also contains subfields with a relatively high percentage of cites given to non-source items, e.g. TROPICAL MEDICINE with 13.5%, REHABILITATION with 15.3%, or MEDICAL INFORMATICS with 20.9%, to mention some medicine-oriented categories. The category with the highest percentage of cites given to non-source items is ENGINEERING, MARINE (56.0%), the one with the lowest percentage is CELL BIOLOGY with 1.6%. The majority of the subfields of the science edition, however, shows relatively low percentages of cites given to non-source items. 73 categories (of 170), including most of the categories with biomedical orientation, show fractions below 10%, 55 are in the range between 10% and 20%, and 23 range between 20% and 30%.

Discussion

Journals are essential for the dissemination of scientific knowledge. The impact factors provided by the Journal Citation Reports (JCR) are predominant measures of a journal's visibility and importance. Therefore, it is advisable to investigate whether there exist other periodicals which are not listed in the JCR but have some weight due to their citedness. The JCR themselves are a source of those other entities because the JCR specify all cited titles including those not being JCR source titles. The data extracted from the JCR 2003 suggest indeed that considerable fractions of cited titles may not be sources of ISI databases. The total number of distinct cited non-source items is above 46,000 for the science edition and nearly 16,000 for the social sciences edition. The effective number of distinct cited non-source entities, however, may be considerably lower because oftenly various forms of the same journal titles exist. This holds true also for the cited titles of journals indexed by ISI (see Christensen, Ingwersen & Wormell, 1996; Stegmann, 1999). So we do not know which fraction of the cited non-source items consists of deviant forms pointing in fact to ISI source titles. From our previous work (e.g. Stegmann, 1999; Stegmann & Grohmann, 2001; Grohmann & Stegmann, 2004) we have several files containing different forms of cited JCR-listed journal titles, and we will use these data to get an impression of the extent to which those non-standard abbreviations of JCR source titles collect cites which are classified, at present, as cites given to non-source items. From preliminary experiments we conclude that up to 30% of the cites collected by non-source items are given to aberrant forms of JCR titles. Yet our study gives evidence that in the social sciences and in some science disciplines notable parts of cited items are not indexed in ISI databases.

The question whether the cites given to titles not listed in the JCR are equally distributed over the journals of a category or concentrated in journals with high or low impact factors was not (yet) investigated but deserves certainly closer attention. The question to which type the referenced material belongs is difficult to answer. The cited character strings in the JCR point to source titles (e.g. journal titles) but not to individual contributions (e.g. journal articles or book chapters). The majority of the cites is given to journals listed in the JCR (see Results); thus we conclude that most of the cited non-source items represent also journals. A rationale for detailed studies - which rely on ISI article databases - concerning the type of cited non-source materials has been proposed by Butler & Visser (2004) and Visser & Moed (2004).

Conclusion

Analysis of the *Citing Journal Listings* of the JCR 2003 reveals that publications not indexed in ISI

databases contribute to some extent (depending on the scientific field) to the citation base of knowledge. Bibliometric studies dealing with citation rates and relying only on ISI databases should always be aware of this fact and, perhaps, document to which extent possibly relevant cited items were not taken into account.

References

- Allee, R.W. (1987). The relationship between primary and secondary publishers: ISI's journal selection guidelines and use of the Journal Citation Report. *Serials Librarian*, 13, 47-55.
- Butler, L. & Visser, M. (2004). Extending citation analysis to non-source items. *Book of Abstracts, Eighth International Conference on Science and Technology Indicators*, Leiden: 55-56.
- Christensen, F.H., Ingwersen, P. & Wormell, I. (1996). Online determination of the journal impact factor and its international properties. *Scientometrics*, 40, 529-540.
- Garfield, E. (1996). How can impact factors be improved? *British Medical Journal*, 313, 411-413.
- Institute for Scientific Information (2003). *Journal Citation Reports on CD-ROM*, science ed., social sciences ed.
- Schoonbaert, D. & Roelants, G. (1996). Citation analysis for measuring the value of scientific publications: Quality assessment tool or comedy of errors? *Tropical Medicine and International Health*, 1: 739-752.
- Sen, B.K., Karanjai, A. & Munshi, U.M. (1989). A method for determining the impact factor of a non-SCI journal. *Journal of Documentation*, 45: 139-141.
- Spaventi, J., Tudor-Šilovic, N., Maricic, S. & Labus, M. (1979). Bibliometrijska analiza znanstvenih casopisa iz Jugoslavije (Bibliometric analysis of scientific journals from Yugoslavia). *Informatologia Yugoslavica*, 11: 11-23.
- Stegmann, J. (1997). How to evaluate journal impact factors. *Nature*, 390: 550.
- Stegmann, J. (1999). Building a list of journals with constructed impact factors. *Journal of Documentation*, 55 (3): 310-324.
- Stegmann, J. & Grohmann, G. (2001). Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in the Journal Citation Reports. *Scientometrics* 50: 483-502.
- Grohmann, G. & Stegmann, J. (2004). German Medical Faculties in the nineties: online bibliometric analysis, *Book of Abstracts, Eighth International Conference on Science and Technology Indicators*, Leiden: 106-107.
- Visser, M. & Moed, H.F. (2004). Measuring the impact of non-ISI source items. *Book of Abstracts, Eighth International Conference on Science and Technology Indicators*, Leiden: 201-202..