Jointly published by Akadémiai Kiadó, Budapest
and Kluwer Academic Publishers, Dordrecht

Scientometrics,
Vol. 53, No. 2 (2002) 229–240

Researchers' and users' perceptions of the relative standing of biomedical papers in different journals

GRANT LEWISON

Bibliometrics Research Group, Department of Information Science, City University, London (England)

Journal citation impact factors, which are frequently used as a surrogate measure of research quality, do not correlate well with UK researchers' subjective views of the relative importance of journals as media for communicating important biomedical research results. The correlation varies with the sub-field: it is almost zero in nursing research but is moderate in more "scientific" sub-fields such as multiple sclerosis research, characterised by many authors per paper and appreciable foreign co-authorship. If research evaluation is to be based on journal-specific indicators, then these must cover different aspects of the process whereby research impacts on other researchers and on healthcare improvement.

Introduction

The use of journal citation impact factors as a means to evaluate research has been practised for many years[1–5] but it has not been without its critics.[6–11] These factors have been calculated by several different methods and the reliability of using the "standard" values has been questioned.[12,13] These are calculated by counting citations in year 2 to papers published in the journal in years 0 and 1 and dividing the number of citations by the number of papers. However the method is anomalous as the numerator is based on all publications but the denominator includes only articles, notes, letters and reviews. An alternative measure of citation impact is based on forward citations, i.e., the mean number of citations in years 0 to n to papers published in year 0. This figure is tabulated in a file Journal Expected Citation Rates by the Institute for Scientific Information (ISI). It is available both for all papers and for all cited papers: the figure for the latter is, of course, much higher than for the former.

Journal citation scores are intended as a surrogate measure of the probable citation score of the individual papers within them and to represent the likely influence of the paper on other researchers. Of course individual papers will depart from this mean, some being more highly cited and some less so, although there is a surprisingly good correlation between actual and predicted citations.[14] Some authors use the ratio between

0138–9130/2002/US $ 15.00

actual and "expected" citations for papers from different labs or countries as a measure of quality,[15–17] but this may be affected by the choice of journals in which the scientists publish. They will appear better if they publish in less prestigious journals because their papers will then stand out more, but this may not be the best way to communicate their findings. An example of the failure of this method of analysis was shown in the study[18] of the National Eye Institute in Bethesda, MD. The institute decided to pursue a policy of communicating its research results in general medical journals with high impact factors in order to attract a wider readership. As a result their papers were more frequently cited but because they were on a specialist topic, they were less cited than the average for those journals. So their actual to expected citations ratio went down.

Of course, in biomedicine, the object of research is not merely to accumulate citations, however gratifying this may be for the individual researchers. It is rather to develop an understanding of the subject so that patients may be given better treatment, or prevented from becoming ill or injured in the first place. Account should therefore be taken of how likely the papers in a given journal are to influence clinical practice when they are being evaluated. For example, *Grant* et al.[19] have shown that papers cited on clinical guidelines tend to be in more clinical journals than the average for biomedical research in the UK and these journals tend to be less frequently cited. On the other hand papers cited on patents tend to be published in high impact basic journals.[20,21] There may, therefore, be several different indicators of the utility of a research portfolio, and the relative weight to be accorded to them will depend on the terms of reference of the evaluation being undertaken.

This paper describes an investigation of the subjective ranking of journals most frequently used within individual sub-fields of biomedicine by UK researchers. The objective was to see how well such rankings correlated with citation scores, in other words, whether citation scores could be relied upon as a guide to the subjective rankings of journals in a sub-field. Attention was deliberately focussed on sub-fields of practical utility and in particular on five of the "professions allied to medicine"; in addition two sub-fields that were regarded as more "scientific" were considered for purposes of comparison. The seven sub-fields were as listed in Table 1, with the numbers of papers for UK researchers for the years 1988-1998, and the names of the experts who defined the subject areas, see below.

There was rather little correlation between the rankings given by the researchers and the citation scores and attention then turned to whether the amount of correlation for the different sub-fields was influenced by the characteristics of the sub-field. It was hypothesised that the more "scientific" sub-fields would have a higher correlation between journal mean rankings and citation scores, and indeed this turned out to be the case.

Table 1. The seven biomedical sub-fields used for the study, the numbers of UK papers from 1988-1998 and the experts who defined the filters

Subfield	*Contraction*	N papers	Defining experts
Dietetics	DIETS	3841	Dr. Angela Madden
Multiple sclerosis	MULSC	819	Dr. Lorna Layward
Nursing research	NURSE	2540	Dr. Anne-Marie Rafferty & Dr. Michael Traynor
Occupational therapy	OCCTH	874	Dr. Irene Ilott & Dr. Elizabeth White
Physiotherapy	PHYSL	864	Ms. Sue Madden
Primary health care	PRIME	17303	Prof. Frank Sullivan
Speech & language therapy	SPLTH	1285	Dr. Paul Carding

A subsidiary objective of the study was to see whether the *Science Citation Index* (SCI) and *Social Sciences Citation Index* (SSCI) of the Institute for Scientific Information (ISI) were adequately representative of British output in some subfields. Some journals might be highly influential even though they are not frequently cited. If so, they should perhaps be taken into account when UK research is being evaluated. This is a problem frequently met when research in "peripheral" or developing countries is being considered: if the ISI doesn't process some national journals which are important locally, should their contents nevertheless be included in any evaluative compilation of research outputs? Again, it appeared that some local journals were highly regarded and that their exclusion from the SCI and SSCI, mainly on grounds of citation impact, would have a major impact on the analysis of UK research outputs in two sub-fields (occupational therapy and physiotherapy).

Methodology

The first step was to identify UK papers (articles, notes and reviews) in the SCI or SSCI in each of the seven sub-fields by means of "filters" that selectively retrieved them on the basis of their being in specialist journals or containing specified title words, or both. A full account of the process of filter development is given elsewhere.[22] (The sub-fields were each defined originally in order to conduct studies of UK research outputs in recent years for members of the Research Outputs Database club.[23]) The bibliographic details of these papers (authors, title, source, addresses) were downloaded from the SCI/SSCI to an MS Excel spreadsheet. For two sub-fields, on advice from the experts, papers were also collected from some additional journals (see Table 6) in order to give

better coverage of the subject. The journals processed, with data input in the same format as that used for the SCI/SSCI, were as shown in Table 2.

The papers were classified by the potential impact category (PIC) of the journal in which they were published. This is a number from 1 (low) to 4 (high) based on the five-year mean citation score of the journal, $C_{0\text{-}4}$, as given by ISI's Journal Expected Citation Rates file.[24] Table 3 shows the division of journals by this means, with examples of journals in each category. However, for some of the sub-fields nearly all the journals used by UK researchers are poorly cited and fall in the lowest category.

Table 2. Additional journals processed in the sub-fields of occupational therapy (OCCTH) and physiotherapy research (PHYTH)

Sub-field	Journals added
OCCTH	***Brit. J. Occup. Ther.,*** *J. Occup. Sci. Australia*
PHYTH	*Manual Ther.,* ***Physiotherapy****, Physiother. Res. Int., Physiother. Theory Pract. /Physiother Pract.*

The journals in **bold** contained the large majority of additional papers – indeed the large majority of all UK papers in their respective sub-fields

Table 3. Classification of journals by potential impact category (PIC) according to their five-year mean citation scores, $C_{0\text{-}4}$

PIC	$C_{0\text{-}4}$ values	Examples
4	Above 20	*Ann. Neurol., J. Neurosci., Lancet*
3	Between 11 and 20	*Brit. J. Cancer, Brit. J. Psychiat., BMJ, Pain*
2	Between 6 and 11	*J. Child Lang., J. Neurol., J. Psychosomat. Res., Rheumatology*
1	Below 6	*Aphasiology, Brit. J. Occup. Ther., Clin. Rehabil., Disabil. Rehabil., J. Adv. Nurs., Phys. Ther.*

From the lists of papers, the names of the leading researchers were identified and matched to their most recent addresses. Individually-addressed letters were prepared inviting them to assign marks to journals listed on an attached sheet. In primary health care, the survey was limited to Scottish researchers. In the sub-field of nursing research, additional letters were prepared by Dr. Rafferty and Dr. Traynor of the London School of Hygiene and Tropical Medicine for despatch to nursing research "customers", i.e., nursing superintendents in large district hospitals. The journals listed were the ones most frequently used by British researchers in each sub-field, typically about 20-30 journals but in nursing research and primary health care many more. Because it is known[25] that the relative weights attached to journals on a subjective basis only vary by

about 4:1 between high-impact and low-impact journals, the researchers were invited to assign marks on a scale of 4 with the interpretations shown in Table 4.

Table 4. Markings given to lists of journals by researcher respondents and their meaning

Mark	Interpretation
4	Excellent journals, essential regular reading, the most important results
3	Very good journals, highly desirable to read, many useful results
2	Good journals, worthwhile to read if time permits, some useful results
1	Journals, not normally of interest, few useful results

The degree of compliance with the marking scale varied. Some respondents were punctilious and marked all the listed journals in the way suggested, with about 10% rated "4", a further 20% rated "3", the next 30% rated "2" and the remaining 40% rated "1". Some were more generous so that their mean score was greater than the expected value of 2.0 but many were unable to rate some or even a majority of the journals and contented themselves with marking a few as "4" or "3" and leaving the rest unmarked. In the analysis, only journals given a mark by a respondent were scored so that some of the journals had more ratings than others. The mean rating for each journal was designated the "Relative Esteem Value" (REV) and its standard error was also calculated (and is shown in the figures as bars).

The sub-field lists differed widely as most of the journals were specialist or at least semi-specialist. Only the *British Medical Journal* (now called the *BMJ*) was common to all seven lists, and it was rated on average 3.7 by (Scottish) researchers in primary health care, but only 1.9 by researchers in multiple sclerosis. The journal *Lancet* was on four of the lists: on two it was more highly rated than *BMJ* (MULSC, PRIME) and in two it was rated lower (DIETS, NURSE). The *Journal of Advanced Nursing* was also on four lists. It was rated a top journal (4) by every single respondent, both researchers and research users, in the sub-field of nursing, but only 1.2 by physiotherapy researchers.

The seven sub-fields were scored for their "scientific" content in several ways. None of them is other than a secondary indicator but because they correlate very well, they seem to provide a consensus on the matter as Table 5 shows.

The only anomalous result is that dietetics has few foreign co-authors, but then dietary research may well be particular to an individual country and of little relevance to others, despite being published in international journals. The ranking of this table also accords with the sources of the papers: for MULSC, PRIME and DIETS it was primarily the SCI and for the four other sub-fields the SSCI, with additional non-SSCI journals being used for PHYTH and OCCTH.

Table 5. Ranking of the seven sub-fields on the basis of four secondary indicators of "scientific content". A = number of authors per paper; D = number of addresses per paper; PIC = potential impact category of journal (4 = high, 1 = low; based on citation scores; see Table 3)

Subfield	Mean A	Mean D	D foreign, %	Mean PIC
MULSC	5.18	2.76	33.6	2.60
PRIME	4.31	2.26	17.1	1.87
DIETS	3.75	2.08	6.9	1.74
SPLTH	2.69	1.76	16.4	1.53
NURSE	2.10	1.55	1.3	1.20
PHYTH	1.98	1.44	7.1	1.07
OCCTH	1.88	1.41	3.8	1.11

Results

Figure 1 shows the comparison between the ratings of 24 journals in dietetics as given by 20 respondents, with their standard errors, and their individual five-year citation scores. The correlation is positive but it is not large, suggesting that citations are only of moderate importance to researchers when they attempt to classify journals in their sub-field.

Figure 2 shows a similar comparison in the sub-field of nursing research. The correlation is almost zero. This sub-field is unusual as it is dominated by a single journal, the *Journal of Advanced Nursing*, in which 45% of all UK nursing papers were published. (Subsequently, a similar one-journal dominance was found in occupational therapy and physiotherapy.) However the ratings of the 8 nursing researchers and the 22 nursing practitioners agreed fairly well, as Figure 3 shows. (The main differences were in the social and ethics journals, which were more highly esteemed by the researchers than the practitioners.) This suggests that there is indeed a consensus in this sub-field on which journals are important but they are not the same ones as those that obtain many citations.

The correlations between relative esteem value (REV) and potential impact category (PIC) are quite similar to those with journal five-year mean citation scores ($C_{0\text{-}4}$ values) except that the journals tend to be bunched as there are only four possible PIC values. Somewhat better correlation is obtained if some account is taken of the effects of the research level (RL) of the journal, but this can be either negative (clinical journals are more highly esteemed) or positive (basic journals are preferred). Figure 4 shows the "best" correlation coefficient (r^2 value) obtained for each sub-field and the "scientific content" of the sub-field, as shown by the mean numbers of authors per paper, A. (Very similar graphs are obtained for the other factors shown in Table 5.)

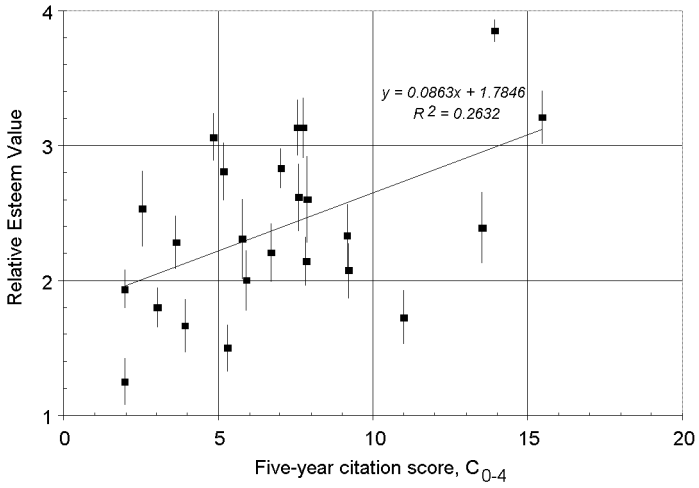

Figure 1. Relative esteem of journals used by UK dietetics researchers (except *Lancet*) compared with their citation scores

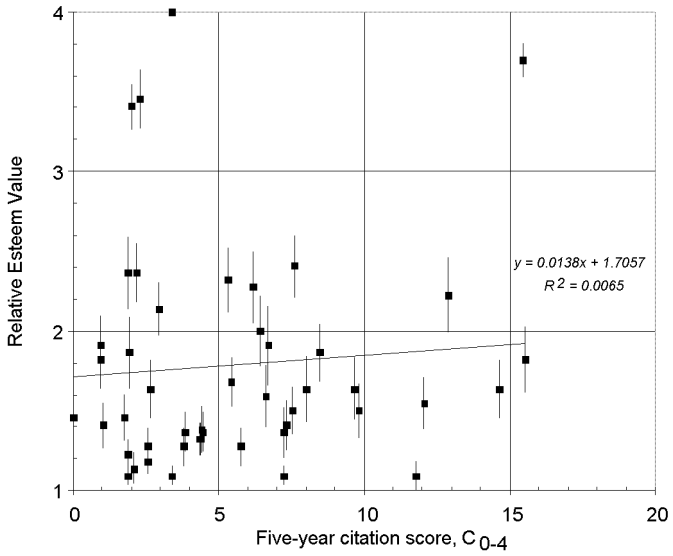

Figure 2. Relative esteem of journals used by UK nursing researchers (except *Lancet*) compared with their citation scores

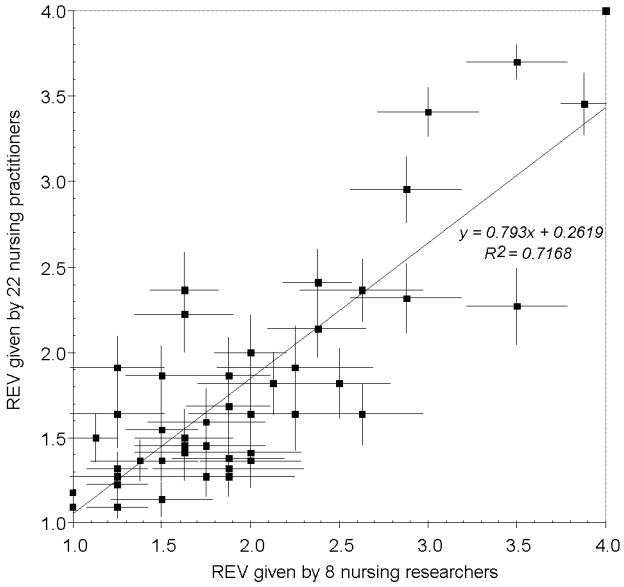

Figure 3. Comparison between relative esteem measures for 48 nursing journals given by researchers and by practitioners

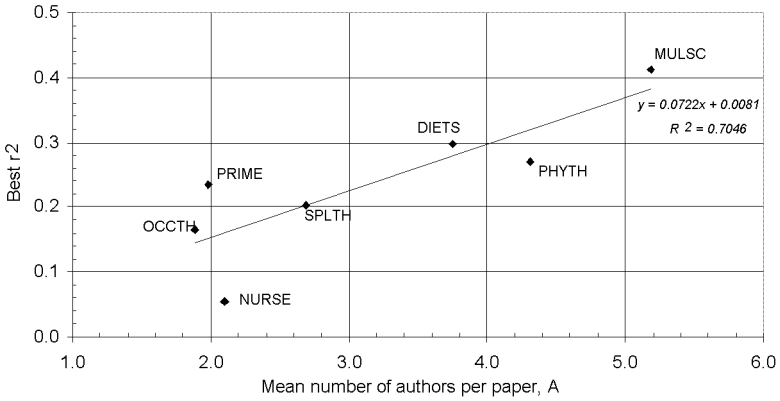

Figure 4. Comparison of best fit correlation with mean number of authors per paper for seven sub-fields

The correlation of the correlations with A is quite good, showing that the more "scientific" sub-fields depend more on citations than the others do in the rating of the popular journals.

What of the additional, non-SSCI, journals covered for the sub-fields of occupational therapy and physiotherapy? Determinations of the numbers of citations received by the UK papers in these journals published from 1994-1996 were made in order to estimate the five-year citation scores (C_{0-4}) of the journals. Of course, since most journals cite papers in the same journal, there were rather few citations found, but the numbers were not zero (except for two journals with 7 or fewer papers). The results are shown in Table 6, with the ratings given by the researchers and their ranking in the sub-fields.

It is clear from this table that at least the two leading journals (*Brit. J. Occupat. Therapy* and *Physiotherapy*) are well esteemed by the researchers, being in the top half of the journals considered; indeed *Physiotherapy* is inferior only to *BMJ* and *Spine*, both well-cited journals.

Table 6. Citation scores of "additional" journals in occupational therapy and physiotherapy and their subjective ratings and rankings

Sub-field	Journal	C_{0-4} value	Rating (1 to 4)	Rank
OCCTH	*Brit. J. Occupat. Therapy*	0.29	2.25	10/22
	J. Occupat. Sci. Australia	0.00	2.00	15/22
PHYTH	*J. Manual. Manip. Therap.*	0.00	2.00	=8/20
	Manual Therapy	0.15	2.80	=3/20
	Physiother. Theor. Pract.	0.35	2.00	=8/20
	Physiother. Res. Int.	0.62	2.00	=8/20
	Physiotherapy	0.18	2.80	=3/20

Discussion

Evaluation of research outputs can be made on the basis of the characteristics of the journals in which the papers are published and increasingly they are being, not least in the UK through the Higher Education Funding Councils' quadrennial Research Assessment Exercise. This is inevitable, given the pressures on evaluators to reach conclusions rapidly before there is time for them to assess the outcomes (including the citation scores) of individual papers. This is not necessarily inappropriate as journals are selective in which papers they accept[26] but it does mean that the characteristics of the journals that are chosen as indicators must be ones that are really relevant to the evaluation. Ideally, multiple partial indicators[27] should be used, and the outputs of

different research groups, institutions or countries compared on a "radar plot" in which the values of the different indicators are shown as markings on scales in different directions. These would include, but not be limited to:

- the citation score of the journals, or a derived indicator such as PIC;
- the subjective Relative Esteem Value of the journals, as described here;
- the frequency with which the journals are cited on clinical guidelines;
- the frequency with which the journals are cited on patents; and perhaps even
- the frequency with which the journals are cited in newspapers.

An example of a hypothetical radar plot for a set of papers, relative to some larger norm group, is shown in Figure 5.

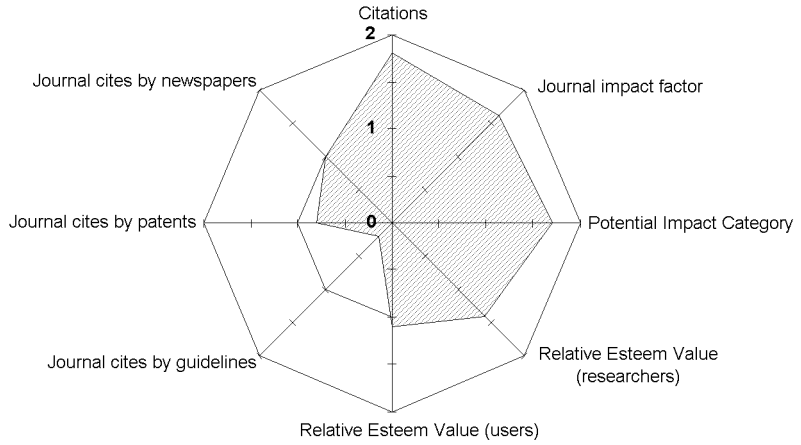

Figure 5. Radar plot of the parameters characterising a selected group of papers relative to some larger norm group and based largely on the properties of the journals in which they are published

The bibliometrics community should concern itself with the construction of some of these other journal indicators in order to give evaluators better tools. It is not likely that these indicators will be universal, except perhaps for the citation of papers by patents where the US Patents and Trademarks Office and the European Patent Office provide detailed data on a world basis. However, since most evaluations are conducted at the national or sub-national level, it would be appropriate to develop national indicators of the relative importance of different biomedical journals to healthcare in individual

countries, based on that country's clinical practices and, perhaps, mass media. It will be important also to confine comparisons to papers (and journals, for the most part) in the same sub-field in order that the evaluation process is seen to be fair.

*

I am grateful to the very many researchers and other respondents who took the time to fill in the journal rating questionnaires. Citations to the papers in some non-SSCI journals were determined by Adriana *Roa Celis*, a visiting PhD student from the University of Campinas, Brazil. The author is a Senior Policy Adviser at the Wellcome Trust, currently on secondment to City University.

References

1. BONITZ, M., Journal ranking by selective impact – new method based on SDI results and journal impact factors, *Scientometrics*, 7 (1985) 471–485.
2. BANOS, J.E., CASANOVAS, L., GUARDIOLA, E., BOSCH, F., Analysis of Spanish biomedical journals by impact factor, *Medicina Clinica*, 99 (1992) 96–99.
3. DAVIS, G., ROYLE, P., A comparison of Australian university output using journal impact factors, *Scientometrics*, 35 (1996) 45–58.
4. GARFIELD, E., Journal impact factor: a brief review, *Canadian Medical Association Journal*, 161 (1999) 979–980.
5. BUTLER, D., France sets high targets for impact factors and patents, *Nature*, 394 (1998) 309.
6. BOOR, M., The citation impact factor – another dubious index of journal quality, *American Psychologist*, 37 (1982) 975–977.
7. EDWARDS, R. G., VAN STEIRTEGHAM, A. C., RICHARDSON, M. J., Impact factors – more misleading than ever, *Human Reproduction*, 8 (1993) 983–985.
8. HANSSON, S., Impact factor as a misleading tool in evaluation of medical journals, *Lancet*, 346 (1995) 906.
9. MOED, H. F., VAN LEEUWEN, TH. N., Impact factors can mislead, *Nature*, 381 (1996) 186.
10. SEGLEN, P. O., Why the impact factor of journals should not be used for evaluating research, *BMJ*, 314 (1997) 498–502.
11. SMITH, R., Unscientific practice flourishes in science, *BMJ*, 316 (1998) 1036.
12. MOED, H. F., VAN LEEUWEN, T. N., REEDIJK, J., A critical analysis of the journal impact factors of Angewandte Chemie and the Journal of the American Chemical Society – inaccuracies in published impact factors based on overall citations only, *Scientometrics*, 47 (1996) 105-116.
13. VAN LEEUWEN, T. N., MOED, H. F., REEDIJK, J., Critical comments on Institute for Scientific Information impact factors – a sample of inorganic molecular chemistry journals, *Journal of Information Science*, 25 (1999) 489–498.
14. LEWISON, G., The quantity and quality of female researchers: a bibliometric study of Iceland, *Scientometrics*, 52 (2001) 29-43. See Figure 7, p. 41.
15. JACOBS, D., INGWERSEN, P., A bibliometric study of the publication patterns in the sciences of South African scholars 1981-96, *Scientometrics*, 47 (2000) 75–93.
16. LUWEL, M., A bibliometric profile of Flemish research in natural, life and technical sciences, *Scientometrics*, 47 (2000) 281–302.

17. NEDERHOF, A. J., VAN WIJK, E., Profiling institutes: identifying high research performance and social relevance in the social and behavioural sciences, *Scientometrics*, 44 (1999) 487–506.
18. CHI RESEARCH INC., private communication.
19. GRANT, J., COTTRELL, R., CLUZEAU, F., FAWCETT, G., Evaluating "payback" on biomedical research from papers cited in clinical guidelines: applied bibliometric study, *BMJ*, 320 (2000) 1107–1111.
20. ANDERSON, J., WILLIAMS, N., SEEMUNGAL, D., NARIN, F., OLIVASTRO, D., Human genetic technology: exploring the links between science and innovation, *Technology Analysis and Strategic Management*, 8 (1996) 135–156.
21. NARIN, F., Patent bibliometrics, *Scientometrics*, 30 (1994) 147–155.
22. LEWISON, G., The definition and calibration of biomedical subfields, *Scientometrics*, 46 (1999) 529–537.
23. DAWSON, G., LUCOCQ, B., COTTRELL, R., LEWISON, G., *Mapping the Landscape – National Biomedical Research Outputs 1988-95*. Wellcome Trust policy report, 9, (1998) See section A3.8, pp. 58–59.
24. LEWISON, G., GRANT, J., JANSEN, P., International gastroenterology research: subject areas, impact and funding, *Gut*, 49 (2001) 295–302. See Figure 2.
25. LEWISON, G., New bibliometric techniques for the evaluation of medical schools, *Scientometrics*, 41 (1998) 5–16.
26. LEWISON, G., VAN ROOYEN, S., Reviewers' and editors' perceptions of submitted manuscripts with different numbers of authors, addresses and funding sources, *Journal of Information Science*, 25 (1999) 509–511.
27. MARTIN, B., The use of multiple indicators in the assessment of basic research, *Scientometrics*, 36 (1996) 343–362.

Received September 30, 2001.

Address for correspondence:
GRANT LEWISON
Bibliometrics Research Group
Department of Information Science, City University
Northampton Square, London EC1V 0HB, England
E-mail: g.lewison@soi.city.ac.uk