*Jointly published by Elsevier Science Ltd, Oxford
and Akadémiai Kiadó, Budapest*

*Scientometrics,
Vol. 41, Nos 1–2 (1998) 17–27*

THE EFFECT OF FUNDING ON THE OUTPUTS OF BIOMEDICAL RESEARCH

G. LEWISON, G. DAWSON

*Unit for Policy Research in Science and Medicine (PRISM), The Wellcome Trust,
210 Euston Road, London NW1 2BE (England)*

(Received August 30, 1997)

The Research Outputs Database (ROD) has been used to investigate the effects of different input variables, including the numbers of funding bodies, on the impact of research papers in a biomedical subfield (gastroenterology). This was determined by the medium-term impact of the journals in which they were published. It was shown that, when account was taken of the effects of the other input factors, the mean impact for a group of papers increased with the number of authors, the type of research (basic more than clinical), and with the number and identity of the funding bodies. However it *decreased* slightly if there were more addresses; whether the paper was multinational had no significant effect. Previous work showing that multi-institution or multi-country papers are more highly cited reached this conclusion because it did not take into account the confounding effect of multiple funding sources, and possibly other factors.

Introduction

Most bibliometric research evaluations make comparisons between the outputs of papers from different institutions or countries. The indicators produced are often contentious (*Collins* 1991) and to be acceptable they must be based on fair comparisons. Indeed, the definition of an appropriate comparison cohort of papers with which those of the organization under evaluation can properly be compared is a difficult task. In some ways it is similar to the selection of a control group of patients to whom a placebo is given during the clinical trial of a new drug: it must be carefully matched to the sample under evaluation with all relevant parameters equalized as far as possible.

Some parameters are now well-recognised as important. For example, citation patterns vary greatly between different fields of science (*Braun* et al. 1995) and even within different areas within a single subfield, so it is important that if citations are to be the basis for the indicator of merit then the control cohort must be in the same

0138–9130/98/US $ 15.00

subfield or subfields as the one being studied. Another influence on citation patterns is the scientific "size" of the country of the author(s), at least in the early years (*Narin* 1994). *Lewison* and *Cunningham* (1989) and *Narin* et al.(1991) have also shown conclusively that international co-authorship leads to more highly cited papers, and this has by now become received wisdom.

This paper is concerned to explore another influence on the output of research papers, namely that of the funders of the research. The hypothesis being advanced is that research supported by several funding bodies, having been considered and selected for funding in competition with many other proposals several times, is likely to be of superior quality to that supported by only a single body or by none. Previously it has not been practicable to examine such a hypothesis because publicly-available databases did not include funding information. [MedLine contains details of funding, but only from the US National Institutes of Health and "other federal government".] However in 1993 the Wellcome Trust began compilation of the Research Outputs Database (ROD), which includes data on acknowledged funding sources for UK biomedical papers in the serial literature, and those implicit from the address field, notably government and commercial laboratories. The design and construction of the ROD were described by *Jeschin* et al. (1995). It is based on the *Science Citation Index* and the *Social Sciences Citation Index* of the Institute for Scientific Information and now covers seven years of publications from 1988 to 1994, with details of over 185 000 papers, so that it is a major resource for analysis purposes.

The ROD has been and is currently being used for several other research projects, investigating:

- the propensity of scientists to acknowledge their funding sources (*Lewison* et al. 1995);
- the sources of support for research in a given subfield (*Higgins* et al. 1996);
- the relationship between funding input and the output of research papers;
- the comparison of outputs from different laboratories in a given subfield (*Lewison* 1998);
- the geographical analysis of the outputs from particular funding bodies (*Lewison* et al. 1997).

The present paper draws on the methods used in these studies and was informed in particular by the results on malaria research reported by *Anderson* et al. (1996) which suggested that papers without a funding body were different in their citation performance from those supported by major funding agencies. A subsequent paper by *Maclean* et al. (1997) showed that citations to malaria papers with 0, 1, 2 and 3 or more funding bodies increased successively with the numbers of funders.

Method

There were two preliminary tasks to perform before the main analysis could be started. The first was to define a subfield "filter" so that the papers retrieved would form a relatively homogeneous set. The procedure for doing this is described elsewhere (*Lewison* 1996) and will not be repeated here except to say that it involves the contribution of one or more experts in the subfield who can both select appropriate title keywords and mark sets of paper titles so that the filter can be successively improved and then calibrated. For the present exercise, the subfield chosen was gastroenterology research, and the filter was designed and checked by three senior UK gastroenterologists. The subfield is typical of biomedical research: it has a good span of levels of research (from clinical to basic); two thirds of the papers are in general rather than in specialist journals so that title keywords are essential to give a good coverage of the subfield; and it is moderate in size, accounting for nearly 7% of UK biomedical research output.

The second task was to develop a system for evaluating the quality, or at least the impact, of the outputs. The conventional measure is to determine the numbers of citations to a group of papers. However, although some citation measures were used, see below, it was not practical for PRISM to determine the numbers of citations to such a large group of papers and in any event some had only been published in 1994 and their citations in 1994-95 would not have given a fair view of their impact. Moreover, because citation number has a severely non-normal distribution, it does not provide a variable which is amenable to the customary statistical analyses, even after transformation (*Anderson* 1989).

It was therefore decided to use journal impact factors as a measure of quality, despite all their limitations (*Moed* and *van Leeuwen* 1996, *Seglen* 1997). In order to avoid one of the criticisms, that they represent short-term influence only, ones based on a five-year citation period were used rather than the standard impact factors which relate numbers of citations in one year to numbers of publications in the journal in the previous two. Each paper in the gastroenterology database was initially given an impact value corresponding to the mean number of citations from 1990-94 to papers published in 1990, as determined in the *Journal Expected Citation Rates* file published by the Institute for Scientific Information, for the journal in which it was published. (New journals had this impact factor extrapolated from values determined for publications in 1991 or later years.) The papers were then listed in descending order of this journal impact value, and the top 10% or decile were given a quality factor, Q, of 10; the next decile, Q = 9, and so on down to the bottom decile which were labelled Q = 1. For a given group of papers, the mean value of their impact decile, Q, therefore gave an indicator of their impact: the norm value was 5.5.

Several different input variables that could have influenced the outcome of the research, as measured by the journal impact decile, were considered. They included the number of authors on the paper, A; the number of addresses, D; whether the paper was internationally co-authored; the number of funding bodies, N; the presence or absence of two particular funding bodies (the Wellcome Trust and another one, designated P); and the research level (from clinical to basic). This parameter is assigned to individual journals by CHI Research Inc. (*Narin*, *Pinski* and *Gee* 1976) on the basis of both expert review and patterns of journal-to-journal citation. Four categories are used:

1 = clinical observation;
2 = clinical mix;
3 = clinical investigation;
4 = basic research.

The records were downloaded from the database into a standard spreadsheet (Excel 5.0) which was used to determine the numerical parameters for each record and the yes/no variables (ICA = internationally co-authored, P = supported by a particular funding body, W = Wellcome Trust funded). The statistical evaluation was conducted by means of a non-orthoogonal analysis of variance, with the algorithm GENSTAT.

Results

The data set consisted of 12 925 records, taken from UK papers on gastroenterology published between 1988 and 1994, whose basic parameters were as listed in Table 1. Papers with large numbers of funding bodies, authors or addresses were grouped together to simplify the statistical analysis. The mean number of authors per paper was 3.92 and the mean number of funding bodies on all the papers was 1.08. However if papers with no acknowledgements are excluded, the mean number of funding bodies per paper was 2.01. Some 12% of the papers were supported by funding body P and 7% were funded by the Wellcome Trust (some were funded by both organisations). The mean number of addresses was 1.91, and 13% of the papers were internationally co-authored.

Table 1
Main parameters of the 12 925 UK gastroenterology papers in the ROD

RL	n	N	n	A	n	D	n
n.a.	163	0	5991	1	1150	1	5667
1	3020	1	3472	2	2467	2	4160
2	5111	2	1769	3	2667	3	1986
3	2616	3	895	4	2293	4+	1112
4	2015	4	421	5	1772		
		5	201	6+	2576		
		6+	176				

RL=research level; N=no. of funding bodies; A=no. of authors; D=no. of addresses

Although the main analysis was intended to be in terms of the mean journal impact decile value of a group of papers, some preliminary citation analysis was carried out to verify that the results obtained earlier (citations increasing with numbers of addresses and with numbers of funding bodies) held also for this set of papers. The results are shown in Figs 1 and 2 respectively for samples of papers randomly drawn from those having the given value of the parameter, n = 294 for most samples but fewer for groups of papers with high numbers of funding bodies.

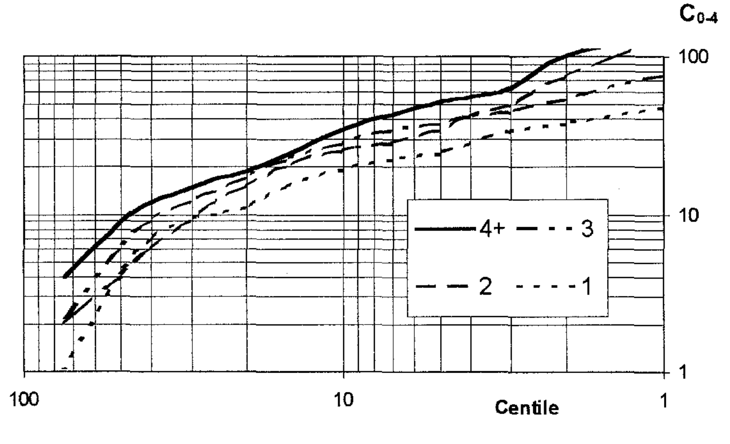

Fig. 1. Citations to UK gastroenterology papers with different numbers of addresses (N=294)

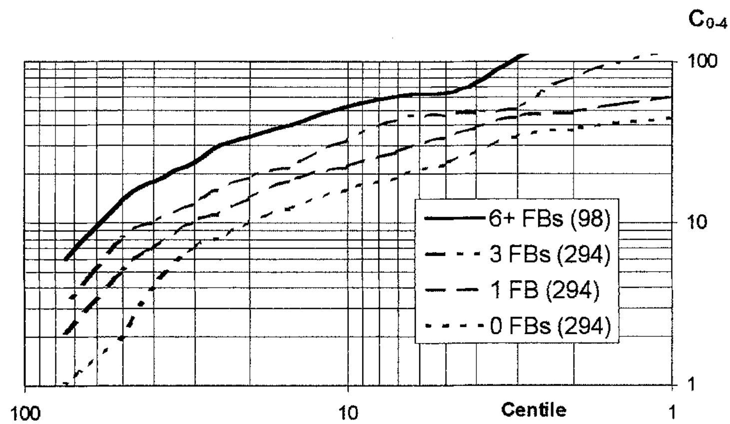

Fig. 2. Citations to UK gastroenterology papers with different numbers of funding bodies

In these figures, the ordinate represents the number of citations in the five years following publication, $C_{0\text{-}4}$, and the abscissa represents the corresponding centile of each group of papers (e.g., the ordinate at the 10th centile of a group of 294 papers would be the number of citations needed for a paper to be in the top-cited 29 papers of that group). It can be seen that the citation scores increase uniformly both with numbers of addresses and with numbers of funding bodies, and that there is a big jump (about 50%) in numbers of citations between papers with no funding body (mostly clinical papers from National Health Service hospitals in this data set) and those with one.

It was necessary to check that this variation in citation performance was closely paralleled by a similar variation in the value of the journal impact decile, Q, in order that reliance could be placed on Q as an indicator of the impact of a group of papers. Since the citation scores shown in Figs 1 and 2 were approximately evenly spaced on a log scale for the different groups of papers, a measure of the mean gap between them for 14 different centiles (1%, 2%, 3%, 4%, 5%, 7%, 10%, 15%, 20%, 25%, 30%, 40%, 50% and 75%) was calculated relative to the average for all the groups, both those with different numbers of addresses (Fig. 1) and different numbers of funding bodies (Fig. 2) and then converted to a citation ratio.

Figure 3 shows the mean Q values for groups of papers with different numbers of addresses, D, shown by open circles, and with different numbers of funding bodies, N, shown by black squares [the citation analysis was also conducted for 2, 4 and 5 funding bodies but the results are not shown in Fig. 2 for the sake of clarity], taken from Fig. 4, plotted on a log scale against the citation ratio as defined above. Although there is some scatter, the correlation is excellent ($r^2 = 0.93$ for both trendlines) and it is remarkable that the two trendlines are virtually parallel: a factor of two difference in the citation ratio between a group of papers and the average for all groups corresponds to a difference in the mean value of Q of 1.32 units.

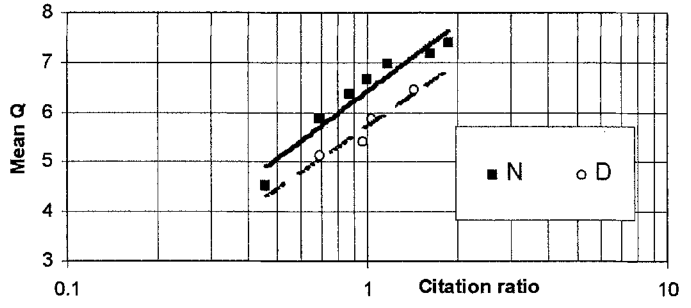

Fig. 3. Comparison of mean citation ratio with mean Q for GASTR papers with different numbers of funding bodies (N) and addresses (D)

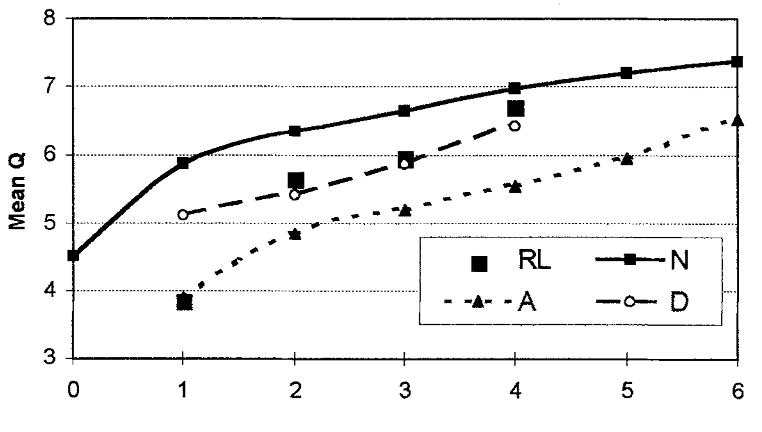

Fig. 4. Variation of journal impact decile with four input factors for UK gastroenterology papers: one-way analysis

Figure 4 shows the variation in the mean journal impact decile, Q, with each of the input variables individually. (The results for research level, RL, are shown as spots rather than as lines as research level is a category rather than a variable with integer values.)

However a rather different picture is shown by Fig. 5, where the mean journal impact decile values for each input factor have been corrected for all the other factors, by means of the non-orthogonal analysis of variance. The variation of mean Q with RL, with A and with N is still positive, though the slopes are less steep; however its variation with D, the number of addresses, is now of negative slope. All the trends are significant at a level below 0.1%.

The effect of international co-authorship, which was positive in isolation (no: Q=5.34; yes: Q=6.25) became slightly negative (no: Q=6.14; yes: Q=6.12) when the other factors (in particular, the larger number of funding bodies associated with internationally co-authored papers) were taken into account. The presence of funding body P and the Wellcome Trust among the acknowledgements on a paper are both positively associated with a higher Q value. After adjustment, Q rises from 5.74 to 6.51 (funding body P) and from 6.02 to 6.23 (Wellcome Trust); the former is significant at the 0.1% level but the latter only at the 5% level.

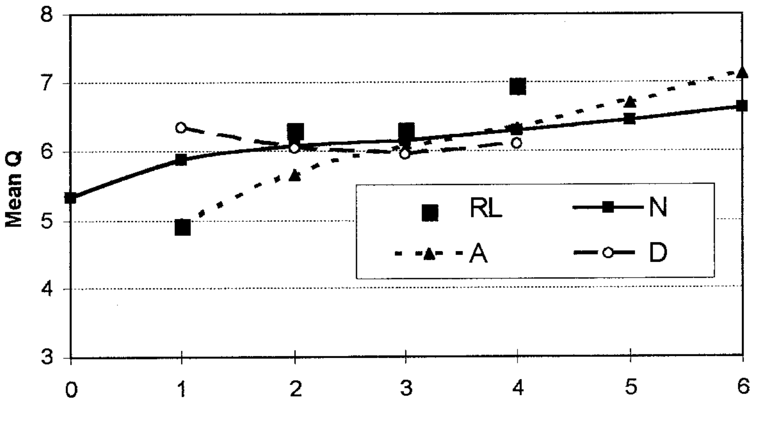

Fig. 5. Variation of journal impact decile with four input factors for UK gastroenterology papers: adjusted analysis

For any given group of papers, there is a distribution of Q values and the analysis leads to a model in which the mean value of Q is predicted. There remains a large variance for the papers within each group, but approximately 65% of the between-group variance has been accounted for by the model.

Discussion

The analysis presented here has shown that the number of authors and the number of funding bodies have a substantial influence on the impact of research outputs. An increase in the number of authors can be seen as an increase in multidisciplinarity, and this is the factor with the biggest potential for an increase in impact of the research output when account is taken of all other input factors. When the number of authors increases from just one to six or more, the mean journal impact decile of the papers rises by more than two, corresponding to a tripling of the number of citations.

The result for the number of funding bodies tends to confirm the original hypothesis that multiple funding sources are correlated with research impact, because their presence in the acknowledgements of a paper show that the work has passed one or more screening processes, mostly through peer review. An increase in the number of funding bodies, especially from zero to one, is correlated with research that is published in journals of higher citation impact and the mean journal impact decile, Q, continues to increase even for six or more funding bodies.

The consequence is that if a funding body seeks to compare the impact of the output of its grantees with the overall value of this parameter for the subfield of interest, it will usually obtain a favourable result, which may engender a false sense of satisfaction if the objective is to isolate the effects of its support. The comparison should then ideally be with a cohort of papers with similar numbers of funding bodies. The Research Outputs Database permits such a comparison: the only other alternative is to look up the papers in a large sample and remove those without acknowledgements before citation numbers are compared, or indicators are produced on the basis of journal impact factors.

The second result, which was unexpected, was that there was no discernable positive effect on journal impact decile of having two or more laboratories co-authoring papers, compared with ones where a similar number of authors, with a similar number of sources of funding, are working on the same type of research in a single laboratory. Indeed, there appears to be a small negative effect, as if co-operation were easier through personal contact than with electronic communications and jet travel. Clearly, it is necessary to check this hypothesis by the examination of additional sets of records. If it is confirmed, it would suggest that the formation of large teams in a single location may be better in terms of the impact of the resulting research papers than the bringing together of smaller groups in different places, such as the formation of "European Laboratories Without Walls", a concept proposed by *Magnien* and *de Nettancourt* (*Magnien* et al. 1989) to promote co-operation in European Community biotechnology research. However ELWWs may be more flexible operationally, and the quality and lasting benefits of research cannot only be measured in terms of publication in high-impact journals – or indeed, by counts of citations.

The model used in the present analysis has accounted for almost two thirds of the (between sub-group) variation in mean journal impact decile for a group of papers; that is variation between papers alike with regard to all of the factors examined. What other factors could be considered to have an influence? One of the suggestions that is sometimes made (*May* 1997) is that the duration of the principal source of funding could be important: there are advocates for both market-driven short-term project grants in universities and the security of long-term core support to tenured researchers in institutes. At present the evidence for the value of different systems is not persuasive. This paper has shown that the confounding effects of the input variables so far considered would certainly have to be taken into account in any investigation. If it were possible to separate out the grantees from a funding body into several groups on the basis of the duration of their funding, their respective outputs could then be compared. Of course, in the absence of data on financial inputs, this analysis would not show which group had been most scientifically productive, but it might at least show which had generated the papers with the potential for the greatest impact.

*

The research described in this paper was supported by the Wellcome Trust as part of its work to promote evidence-based policy making. We are most grateful to Dr Eurof *Walters*, who gave us advice on statistical methodology and carried out the analysis whose results are presented here, and to Mr Robert *Cottrell* of PRISM who undertook the citation analysis.

References

ANDERSON J. (1989), The evaluation of research training. In: D.C. Evered, S Harnett (Eds), *The Evaluation of Scientific Research,* 93-119, Wiley, Chichester.

ANDERSON J., MACLEAN M., DAVIES C. (1996), *Malaria Research: An Audit of International Activity,* PRISM report no 7 (ISBN 1 869835 68 9), The Wellcome Trust, London.

BRAUN T., GLÄNZEL W., GRUPP H. (1995), The scientometric weight of 50 nations in 27 science areas, 1989-93. Part II Life Sciences, *Scientometrics,* 34: 207-237.

COLLINS P.M.D. (1991), *Quantitative Assessment of Departmental Research,* SEPSU Policy Study no 5, The Royal Society, London.

HIGGINS C. et al. (1996), *BBSRC Review of Structural Biology,* Biotechnology and Biological Sciences Research Council, Swindon.

JESCHIN D., LEWISON G., ANDERSON J. (1995), A bibliometric database for tracking acknowledgements of research funding, *Proceedings of the Fifth International Conference of the International Society for Scientometrics and Informetrics*, 235-244. Learned Information Inc, Medford NJ.

LEWISON G., CUNNINGHAM P. (1989), The use of bibliometrics in the evaluation of Community biotechnology research programmes, *Select Proceedings of the First International Workshop on Science and Technology Indicators*, 99-114. DSWO Press, Leiden.

LEWISON G., DAWSON G., ANDERSON J. (1995), The behaviour of biomedical scientific authors in acknowledging their funding sources, *Proceedings of the Fifth International Conference of the International Society for Scientometrics and Informetrics*, 255-264. Learned Information Inc, Medford NJ.

LEWISON G. (1996), The definition of biomedical research subfields with title keywords and application to the analysis of research outputs, *Research Evaluation*, 6: 25-36.

LEWISON G., DAWSON G., ANDERSON J. (1997), Support for UK biomedical research from tobacco industry. *The Lancet,* 349: 778.

LEWISON G. (1998), New bibliometric techniques for the evaluation of medical schools, *Scientometrics,* 41: 5-16.

MACLEAN M., DAVIES C., LEWISON G., ANDERSON J. (1997), Evaluating research activity and impact at the level of the funding body, submitted to *Research Evaluation,* in press.

MAGNIEN E., AGUILAR A., WRAGG P., DE NETTANCOURT D. (1989), Laboratoires Européens sans murs (European Laboratories Without Walls). *Biofutur*, 17 November, 17-30.

MAY R. (1997), The scientific wealth of nations, *Science,* 275: 793-796.

MOED H. F., VAN LEEUWEN T.N. (1996), Impact factors can mislead, *Nature,* 381: 186.

NARIN F., PINSKI G., GEE H.H. (1976), Structure of the biomedical literature, *Journal of the American Society for Information Science* 27: 25-45.

NARIN F., STEVENS K., WHITLOW E.S. (1991), Scientific co-operation in Europe and the citation of multinationally authored papers, *Scientometrics,* 21: 313-324.

NARIN F. (1994), Patent bibliometrics, *Scientometrics,* 30: 147-155.

SEGLEN P.O. (1997), Why the impact factor of journals should not be used for evaluating research, *BMJ,* 314: 498-502.