Scientometrics, Vol. 14, Nos 3–4 (1988) 213–233

BIBLIOMETRIC PROFILES FOR BRITISH ACADEMIC INSTITUTIONS: AN EXPERIMENT TO DEVELOP RESEARCH OUTPUT INDICATORS

M. P. CARPENTER,** F. GIBB,*** M. HARRIS,* J. IRVINE,****
B. R. MARTIN,**** F. NARIN**+

**Advisory Board for the Research Councils, Elizabeth House, York Road, London SE1 7PH (UK)*

***CHI Research/Computer Horizons Inc., 10 White Horse Pike, Haddon Heights, New Jersey 08035 (USA)*

****Department of Information Science, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH (UK)*

*****Science Policy Research Unit, University of Sussex, Falmer, Brighton BN1 9RF (UK)*

(Received January 21, 1988)

In this paper, we report the results of an exploratory study commissioned by the Advisory Board for the Research Councils to produce bibliometric research profiles for academic and related institutions within the UK. The approach adopted is based on the methodology developed by CHI Research whereby publications from a given institution are weighted according to the influence of the journal in which they appear. Although certain technical limitations were encountered with the approach, the study nonetheless yielded potentially useful information on the comparative research output of British universities and polytechnics.

Introduction: background to the study

In May 1986, the University Grants Committee (UGC)[1] sent letters to all British universities reporting the results of their assessment of the research strength of academi departments. This exercise provoked a storm of controversy,[2] not just among those universities and departments faring badly in the assessment. That the UGC was led to expose itself in this way to the fury of a significant section of the academic community reflects several factors. First, and probably most important, are the severe financial constraints imposed on universities by a Conservative government wedded to reducing public expenditure[3] and to obtaining greater public accountability and better 'value for money' from its investments in R&D. At the same time, the costs of con-

+The findings and conclusions presented are those of the authors alone and do not necessarily represent the views of their institutions or the Advisory Board for the Research Councils. Correspondence concerning the paper should be addressed to *Martin* at SPRU.

ducting frontier research have been escalating rapidly in many areas as equipment has grown more sophisticated. Last, it has become clear over recent years that much academic research is likely to prove strategically important in terms of providing the knowledge base for a range of new technologies and ultimately new industries.[4] Since Britain cannot afford to compete in all areas of research, the only realistic option is to pursue policies for science founded on greater selectivity and concentration, choosing to support frontier research in only some areas, and, within these, focussing resources on institutions likely to make the greatest contributions. It was an attempt by the UGC to identify the stronger research groupings in each field that led it to undertake for the first time a comprehensive ranking of university departments.

The methodology adopted in the UGC ranking exercise was relatively simple. In 1985, all universities were asked to prepare for each of their 'cost centres'[5] a short two-page profile of research activities together with a list of five important publications produced over the previous five years. The submissions were then considered by discipline-based UGC subcommittees which also took into account other information such as the value of research grants from Research Councils and foundations, numbers of research studentships, fellowships and 'new blood' lectureships, and income from industry for commissioned research. Initially a five-point scale was used to rate the research standing of cost centres, but the bottom two categories were subsequently conflated to leave just four – 'outstanding', 'above average', 'average' and 'below average'.[6] The results suggested that research excellence is very unevenly distributed among British universities: at one extreme, Cambridge and Oxford each had over 30 departments rated 'outstanding' and only one 'below average'; at the other were several universities with only one or no 'outstanding' department and large numbers 'below average'.[7]

Given that the UGC rankings were to be used in determining the distribution of part of the Committee's funds to universities, it was probably inevitable that they should become subject to bitter criticism. One objection was that it was unfair to focus on 'cost centres' rather than individual departments, since in many cases this meant that several departments had to share a common list of five key publications. Related to this was the criticism that the ratings failed to take account of the differing sizes of departments and were inherently unfair to smaller ones. Others argued that the criteria employed in the rankings were biassed against certain universities, in particular those with a technological orientation striving to achieve closer links with industry and therefore placing less emphasis on producing academic publications.[8]

Finally, misgivings were expressed about the validity of attempts to assess the research strength of departments on the basis of a sample of just five publications.[9] Instead, it was argued, comprehensive data on published output, and better still on its impact on the scientific community as well, should have been compiled. Indeed, such was the dissatisfaction with the UGC rankings in certain fields that attempts have since been made to compile more representative publication data.[10] One recent study went as far as to conclude that "the UGC ratings have approximately zero vilidity."[11]

Besides the efforts of UGC to identify comparative research strengths and weaknesses, other UK agencies have begun to explore the possible use of quantitative output indicators in determining their future funding policy. In particular, the Agricultural and Food Research Council and the Natural Environment Research Council have been experimenting with the use of bibliometric data as a means of monitoring and evaluating the research institutes and projects they support.[12] The Economic and Social Research Council has also started to examine the utility of bibliometric indicators for assessing the performance of social scientists.[13] However, in these bodies (unlike with the UGC) such indicators have generally not yet begun to influence policy-making directly.

Among those who have attempted to construct output or performance indicators for fundamental research, there is near universal agreement on the need for data not just on numbers of publications, but also on their subsequent scientific impact.[14] Analysts have usually attempted to gauge the latter by counting the number of times publications produced by a particular research group or institution have been cited in subsequent publications.[15] There are, however, three problems with this approach. First, it can prove costly and labour-intensive to match each and every citation to a given publication, particularly if one tries to include references containing misprints (e.g. misspelt authors names and incorrect volume or page numbers). Second, there is a 'time-lag' effect in that it typically takes two or three years to establish whether a publication is found useful and cited by others. Third, since referencing habits vary considerably across scientific fields, citation counts cannot be employed to make direct comparisons across fields without first 'normalizing' the figures to take account of such variations. In an attempt to overcome these problems, *Narin* and his colleagues at CHI Research have over recent years developed the influence weight procedure.[16]

The influence weight procedure

The basis of the CHI method is that, rather than counting citations to a set of papers individually, one instead uses the 'influence' of the journals in which they are published as a surrogate measure of their impact. Given that the actual impact of papers within a given journal will vary widely, such an approach will clearly only yield reliable results for larger aggregations of papers (or researchers). The question of how many papers per institution are required before significance can be attached to such results is addressed below.

As with much bibliometric work, the starting point for constructing research profiles and influence indicators is the *Science Citation Index* (*SCI*) produced annually by the Institute for Scientific Information (ISI). The first step carried out by CHI was to classify the 3000 journals[17] included in the *SCI* into eight[18] fields (mathematics, physics, chemistry, earth and space science, engineering and technology, clinical medicine, biomedical research and biology) and approximately 100 subfields. Here, the main problem concerns multidisciplinary journals such as *Nature,* where some 100 journals have had to be categorized according to the estimated proportion of articles they contain from each subfield. Using this classification scheme,[19] one can determine how many papers have been published by a given institution in each field and subfield, and also generate a number of related indicators such as the 'activity index' (this is discussed below).

The next step was to calculate for each journal the average influence per paper. This is defined as the average number of times papers in the journal are cited, each citation having first been weighted by the 'influence weight' of the citing journal – in other words, a citation from an influential journal is scored more highly than one from an obscure journal. The 'influence weight' of a journal is in turn defined as the weighted number of citations it receives from other journals, *normalized by the number of references it gives.* As can be seen, this is similar to the definition of average influence per paper, but it takes into account the fact that papers in some fields (e.g. biochemistry) contain more references on average than those in others (e.g. mathematics). It is also apparent that the definition is partly circular – one does not know how heavily to weight the citations from particular journals in advance of calculating their influence weights. One solution, therefore, is to adopt an iterative approach, starting with an initial arbitrary set of journal influence weights, performing the first set of calculations, reinserting the new weights into the second set, and repeating the approximation process until the resultant weights settle down.[20] In this study, journal influence weights were calculated on the basis of citations listed in the 1982 edition of the *SCI.*[21]

The final step in the influence methodology was to generate publication lists for all relevant UK research institutions and, using the journal influence data, calculate for each the average influence per paper for the various fields and subfields. The resulting 'influence per paper' figure for each field or subfield could then be compared with the average value for a given population of institutions. Since figures for 'influence per paper' vary widely across fields, it is convenient to express the results for each institution in terms of an 'influence score' indicator – that is, the number of standard deviations above or below the mean of a given institution's influence per paper figure compared with the UK average.

The ABRC study[22]

The overall aim of the project was to construct research profiles for all UK universities and medical schools, polytechnics (and the equivalent Scottish institutions), research council establishments, and most government laboratories engaged in civil R&D. However, we shall concentrate here on the profiles for universities and polytechnics resulting from research published in 1983–1984.[23]

The project represented the most recent in a series of science policy studies supported by ABRC. The first was conducted in 1983–84 and was concerned with the testing in a policy context of bibliometric indicators of research activity in five selected specialties. Three approaches were explored: the co-citation technique pioneered by ISI and the Center for Research Planning; the co-word methodology developed by *Turner* and his colleagues in Paris; and the publication and citation-based 'partial indicators' of scientific output employed by *Martin* and *Irvine* at the Science Policy Research Unit (SPRU). Details of the study's results can be found elsewhere.[24]

A year leter, ABRC funded a study by the Science and Engineering Policy Studies Unit (SEPSU) of the Royal Society and Fellowship of Engineering. This adopted a two-fold approach to assessing the health of British science. First, in order to obtain international data on Britain's relative scientific performance, SEPSU utilized the 'Science Literature Indicators Data-Base' constructed by CHI Research for the US National Science Foundation. Second, for more disaggregated information relating to individual specialties of solid-state physics and genetics, the bibliometric techniques developed at SPRU were employed. Again, the results can be found elsewhere.[25]

These two earlier ABRC studies were somewhat specialized in scope. The first, although examining both national and institutional performance, dealt with only five research fields (for example, protein crystallography). The second, while it attempted to encompass a much broader sweep of science, concentrated on national performance, and yielded data on institutional output for only a few narrow specialties (e.g. spin glass).[26] In 1986, therefore, ABRC commissioned the present study to provide a more global overview of the research performance of UK universities and other research institutions across the broad range of science and engineering. The underlying objective was to test whether reliable research output indicators could be produced which might help the Research Councils and other funding agencies in determining future policies.

There were three main components to the study. The first, which was carried out by *Gibb* and his colleagues at the University of Strathclyde, involved analyzing the 1983 and 1984 *SCI Corporate Index* to develop a directory of all relevant UK institutions to be included in the study (499 in total).[27] Since authors sometimes use differing addresses for the same institute or department, it was necessary to compile a comprehensive thesaurus of all address variants for each of the 499 institutions. Where problems were encountered in deciding whether to include an organization, standard reference books on British universities and research institutions[28] were consulted. In some cases, telephone calls were also made to seek clarification on unresolved points. Among the difficulties arising that should be noted were those involved in:

(1) separating certain elements attributable to UMIST from those that are part of the Victoria University of Manchester, and allocating elements with a general affiliation to the University of London to specific colleges;

(2) identifying the teaching hospitals associated with medical schools, and separating the output of medical schools from that of their parent universities;

(3) distinguishing research council units situated on a university campus from the host university.[29]

Using the thesaurus of institutional addresses,[30] *Gibb* and his co-workers identified from the 1983 and 1984 *SCI Corporate Index* tapes all papers published by the 499 institutions, each paper being given an institutional code. These data were used by *Carpenter* and *Narin* of CHI Research to conduct the second component of the study. First, from the *SCI Corporate Index* tapes, they were able to identify all co-authored papers and to allocate credit fractionally to the institutions involved.[31] For each institution, the papers were then sorted into fields and subfields on the basis of the journals in which they were published. This is important to bear in mind when one comes to examine the eventual results since it means that the profile data, say for physics in a particular university, relate *not* to papers published by members of the physics *department,* but to articles by university staff appearing in physics *journals*. Last, CHI compared the publication and influence data for each organization with those for other British institutions in order to calculate, for example, what percentage of national effort in each subfield they accounted for, and how many standard deviations their influence per paper was above or below the mean either for other similar institutions or for all 499 UK organizations (i.e. the 'influence score'). Further details of the various bibliometric indicators constructed are given below.

The third component was conducted by *Irvine*[32] and *Martin* of the Science Policy Research Unit (SPRU). Besides acting as consultants to ABRC for the overall study and undertaking analysis of the final results, they were responsible for two other tasks. The first was to develop user-friendly software for processing the large volume of basic bibliometric data supplied by CHI.[33] This was undertaken by a specialist programmer, and a user-manual prepared.[34] The other task carried out by SPRU involved an attempt to establish empirically the number of papers per institution required to yield reliable results for the influence indicators. Statistical analysis by CHI suggested that the number of papers lay in the range between 20 and 50, but there were some technical doubts about the accuracy of the method employed to derive these parameters. In addition, the high level required for significance posed problems for the UK data since, for many of the institutions, the number of papers per field lay below this threshold. The approach adopted to establish this threshold empirically involved selecting a sample of institutions, carrying out a full citation analysis of their papers, and correlating the resulting citation statistics with the corresponding influence data. The initial step involved selecting five universities which published around 25 papers in 1983–84 in physics and five publishing a similar number in biology, in both cases covering as wide a range of 'influence per paper' as possible. Citations to these papers in 1984, 1985 and 1986 were counted manually using the published volumes of the *SCI.*[35] For each university, the total number of citations and average number of citations per paper were calculated, and the results correlated with those for total influence[36] and influence per paper. The resultant figures are shown in Table 1 below. Overall, the correlation was rather lower than expected – for total influence and citations it was 0.32, and for influence per paper and citations per paper 0.53.[37]

The exercise was then repeated for 15 universities publishing 45–50 papers in 1983–84, again in either biology or physics. The results are given in Table 2. This time the two correlations were respectively 0.80 and 0.84.[38] One interpretation of the results from this exercise, therefore, is that the influence indicators are unlikely to be reliable for institutions producing less than 30–40 papers per field (or subfield). However, one could alternatively argue that the influence data represent a *better* indicator than simple citation counts because they involve giving greater weight to citations from more influential journals. Having discussed details of the methodology employed, let us now turn to the results.

Illustrative results from the study

A specimen set of results for one institution, University A, is shown in Table 3.[39] In order to understand the meaning of each indicator, let us consider the top row which gives figures for biology. From the first column, one can see that University A published 46.5 papers in biology journals (after fractionating collaborative articles). Since publication practices vary considerably between fields, this figure on its own gives little indication of whether, in relation to its total output, the university is publishing an above or below average number of biology papers. The 'activity index'

Table 1
Comparison of influence and citation data for universities publishing around 25 papers in 1983–84

	CHI data			Citation data		
	No of papers[a]	Total influence	Influence per paper[b]	No of papers[c]	Total citations	Citations per paper
(1) BIOLOGY						
University 1	21.5	354	22.3	33.4	135	4.04
2	29.8	474	19.1	35.5	69	1.94
3	29.8	477	16.7	42.0	177	4.21
4	18.4	235	14.4	20.1	50	2.49
5	20.8	190	9.4	26.0	32	1.23
(2) PHYSICS						
University 6	30.3	637	21.1	35.6	102	2.87
7	19.0	270	16.9	35.0	194	5.54
8	20.4	287	14.2	24.1	44	1.83
9	31.5	412	13.5	35.2	96	2.73
10	28.9	288	10.5	30.6	57	1.87

[a]After fractionating collaborative papers.
[b]After excluding papers published in journals for which CHI has not calculated an influence weight.
[c]Without fractionating collaborative papers, but using the same procedure as CHI for treating publications in journals covering two or more fields.

is a way of normalizing the 'raw' publication count; it is defined as follows:

$$\text{Activity index} = \frac{\%\text{ of University A's papers in a given field}}{\%\text{ of all universities' papers in that field}}$$

For all UK universities, 12.7% of papers were in biology journals, while for University A the corresponding figure was 9.7% (see the third column). The activity index is therefore 9.7/12.7 or 0.8 (second column), i.e. University A published fewer biology papers in relation to its total output than average, accounting for 1.5% of the total for UK universities (fourth column).

The fifth column of Table 3 relates to the fact that, for a few comparatively small or new journals, CHI was unable to calculate a meaningful 'influence weight' figure.[40] However, 85.8% of University A's biology papers were in journals for which an influence weight was calculated, and these had an average influence per paper of

Table 2
Comparison of influence and citation data for universities publishing around 45–50 papers in 1983–84

		CHI data			Citation data		
		No of papers[a]	Total influence	Influence per paper[b]	No of papers[c]	Total citations	Citations per paper
(1) BIOLOGY							
University	11	46.1	935	24.0	49.1	160	3.26
	12	39.6	782	22.6	44.3	155	3.50
	13	48.7	865	19.0	70.7	297	4.20
	14	57.9	957	17.5	69.5	253	3.64
	15	48.9	735	16.0	51.6	157	3.04
	16	47.6	663	14.8	61.9	133	2.15
	17	52.4	662	13.9	75.7	138	1.82
	18	42.0	273	9.5	59.0	66	1.12
(2) PHYSICS							
University	19	44.3	1508	34.1	54.6	308	5.64
	20[d]	43.3	1013	24.5	55.6	191	3.44
	21	39.1	857	23.1	47.6	248	5.21
	22	52.9	908	19.8	54.8	215	3.92
	23	49.5	718	17.5	55.5	235	4.23
	24	53.5	670	14.1	60.1	118	1.96
	25	54.2	408	8.7	59.9	101	1.69

[a]After fractionating collaborative papers.
[b]After excluding papers published in journals for which CHI has not calculated an influence weight.
[c]Without fractionating collaborative papers, but using the same procedure as CHI for treating publications in journals covering two or more fields.
[d]Based on a random sample of two-thirds of physics papers produced by University 20.

23.6. This compares favourably with the figure of 17.5 for all biology papers produced by UK universities. University A's average influence corresponds to 2.6 standard deviations above the UK mean, so it has an 'influence score' of 2.6 (see seventh column). The final column, 'average research level', describes how basic or applied are the journals in which University A published. CHI has classified the 3000 journals scanned by *SCI* into four categories, with level 1 corresponding to journals of an applied technological nature, and level 4 representing very basic scientific literature.[41] A research level of 3.6 means that most of the biological research by University A is relatively basic, while, as one might expect, research in the field of engineering is more applied.

Table 3
Specimen bibliometric profile for University A, 1983–84

	Number of papers	Activity index	% Internal effort	% Total UK effort	% papers with influence	Average influence	Influence score	Average research level
Biology	46.5	0.8	9.7	1.5	85.8	23.6	2.6	3.6
Biomedical research	89.8	1.2	18.8	2.5	93.9	63.3	2.7	3.9
Chemistry	122.4	1.3	25.6	2.5	98.1	26.9	2.2	3.9
Clinical medicine	21.3	0.5	4.4	0.9	78.8	16.0	0.3	2.9
Earth & space science	40.4	1.1	8.4	2.1	97.3	26.1	0.2	4.0
Engineering	26.0	0.5	5.4	1.0	76.1	5.9	–1.1	2.0
Mathematics	11.0	0.5	2.3	1.1	95.5	4.0	–1.2	3.3
Physics	121.3	1.4	25.3	2.8	93.8	21.4	0.9	3.7
All fields combined	478.7	1.0	100.0	2.0	92.8	30.1	5.1	3.7

From other rows in Table 3, one can see that University A published relatively heavily in physics (with an activity index of 1.4), chemistry (1.3) and biomedical research (1.2), the first two fields each accounting for a quarter of its total published output. Furthermore, in all three fields, University A used journals of an above average influence, recording (positive) influence scores of 0.9, 2.2 and 2.7 respectively. In engineering, mathematics and clinical medicine, in contrast, comparatively few papers were produced, and in the first two areas negative influence scores were obtained.[42] Overall, University A's research output tends to be concentrated in fields where journals obtain relatively high values for 'average influence per paper' (i.e. biomedical research, chemistry and physics); less effort is devoted to fields like engineering and mathematics where low journal influences are the norm. This is one reason why, for all fields combined (see the bottom row of Table 3), the university obtained an influence score of 5.1, one of the highest for all British institutions. Similar profiles can be produced from the data set for other UK universities, polytechnics, research council establishments, and certain government R&D laboratories. In addition, the eight fields can be further disaggregated into 99 subfields. However, the number of papers per institution is then in most cases less than 20, with the result that the influence data are generally not significant.

Since the results for individual institutions have yet to be made public, let us instead look at how publications and influence are distributed across the UK university sector. Table 4 summarizes the results for four fields. Excluding those without departments of biology,[43] the top five universities (i.e. the top decile) published 25.7% of all papers in biology journals and the top 12 (the top quartile) 48.8%. Conversely, the bottom quartile and decile produced only 4.6% and 1.1% of publications respectively. The figures in the second column show that, in terms of total influence, there is a slightly higher degree of concentration in the leading universities, with the top five, for example, accounting for 26.8% of influence compared with 25.7% of papers. As one moves across the table through chemistry and engineering to physics, one finds the leading institutions in each field obtaining an increasing share of publications and influence. In physics, where the costs of carrying out first-rate research are greater and 'critical mass' effect might therefore be expected to be more pronounced, the top five published no less than 38.2% of papers which accounted for 43.1% of influence, while for the bottom 27 universities with physics departments the corresponding figures were only about half this (20.1% and 19.0% respectively).

Another important issue upon which the results of the study can shed light is the research performance of polytechnics. For many years there has been considerable debate about whether staff at polytechnics should be expected to carry out research in addition to their teaching duties. In many cases, lack of time and resources greatly restricts the research ambitions of polytechnic lecturers. Even so, the figures in

Table 5 for two fields suggest that the research output of leading polytechnics is at least as good as that of the lowest placed universities. In chemistry, there is little difference in numbers of publications and influence[44] between the bottom five universities and the top five polytechnics. In biology, in contrast, the output of the leading polytechnics (including equivalent Scottish institutions) is appreciably better than that of the bottom five universities with biology departments. Overall, however, the research profiles of polytechnics are on average notably inferior to those of universities.

Finally, let us consider how the results of this study compare with the UGC rankings of university departments discussed earlier. There are, however, two main problems in attempting to draw such comparisons. First, as mentioned earlier, the breakdown by field for the profile data is based on *journal subject* classification rather than the *departmental affiliation* of authors. Second, many of the UGC 'cost centres' correspond only loosely with the CHI field categories.[45] Nonetheless, two fields where there is a reasonably good correspondence are chemistry and physics, and we shall therefore concentrate on these.

Table 6 gives research profile data for the 54 UK universities with physics departments, together with their respective UGC ratings. The universities have been ranked in terms of the total influence of their physics publications. It is significant that the four universities with the greatest influence were the only four to be judged 'outstanding' by UGC, while four of the next five were classified as 'above average'. At the other extreme, 10 of the bottom 11 universities with least influence in physics received a 'below average' rating. Even so, there are several institutions for which the UGC ranking bears little obvious relationship to the bibliometric profile data. For example, Universities 39 and 46 received higher rankings than their publication and influence records might indicate, while Universities 6 and 12 seem to have been comparatively harshly treated. Overall, however, there is a correlation of 0.63 between total influence and UGC ranking. Interestingly, the correlation between numbers of papers and UGC rating is at least as high (0.65), which perhaps suggests that the UGC assessments may in some cases have been influenced by the total output of physics departments than by their 'productivity' – that is, their output adjusted for their size or 'inputs'. This is apparently borne out by the small correlations between the UGC rankings and the size-independent indicators of average influence per paper and influence score (the correlations for these were 0.22 and 0.34 respectively).[46]

Although the data are not reproduced here, a similar picture emerges in relation to chemistry. Again, the five universities with the greatest total influence were the only ones to be judged 'outstanding' by the UGC, while the bottom eight were all rated 'below average'. The correlation between total influence and the UGC rankings was indeed rather higher in chemistry at 0.77, as was that between numbers of publica-

Table 4
Distribution of papers and influence across UK university sector, 1983–84

Universities	Biology (n = 50)		Chemistry (n = 54)		Engineering (n = 48)		Physics (n = 54)	
	% of papers	% of influence	% of papers	% of influence	% of papers	% of influence	% of papers	% of influence
Top 10%	25.7	26.8	26.8	28.8	31.2	37.9	38.2	43.1
25%	48.8	49.0	48.4	51.5	53.8	59.4	57.8	62.6
50%	77.2	77.4	75.8	77.8	79.0	82.2	79.9	81.0
75%	95.4	96.4	91.7	92.8	93.0	94.4	93.4	94.2
90%	98.9	99.0	97.7	98.1	98.1	98.4	98.4	98.5
100%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

tions and UGC rating (0.76).[47] Again, however, there were a few institutions which appear on the basis of the profile data to have been treated either particularly generously or unduly harshly.

Conclusions and policy implications

The study reported here should be viewed as no more than an experiment to investigate the validity and potential policy utility of bibliometric research profile data.[48] As observed earlier, the methodology possesses several advantages over other approaches to constructing output indicators for science. In particular, it yields comprehensive data, in this case for several hundred UK research institutions, at relatively low cost. Furthermore, compared with traditional citation analysis, there is less of a time-lag between the calculation of the profile statistics and the research activity to which they relate.

Given that this was a pioneering exercise, however, it is perhaps inevitable that appreciable methodological problems were encountered. First, the categorization of

Table 5
Publication and influence data for leading universities and polytechnics and for the lowest placed universities, 1983–84

	Biology			Chemistry	
	No of papers	Total influence		No of papers	Total influence
University 1	160.2	1800	University 1	330.5	8316
2	150.9	2869	2	325.8	8309
3	147.3	3165	3	227.2	4777
4	136.0	2087	4	214.2	4929
5	130.2	1620	5	165.5	3398
46	9.7	131	50	27.7	601
47	8.9	122	51	22.4	473
48	5.6	55	52	21.8	193
49	4.1	60	53	19.5	308
50	3.5	58	54	18.3	414
Polytechnic 1	31.5	372	Polytechnic 1	27.0	591
2	27.4	271	2	26.6	474
3	25.3	216	3	22.7	41[a]
4	16.5	236	4	16.4	100
5	15.3	144	5	12.0	250

[a]Only 26% of the chemistry papers from the third placed polytechnic were published in journals for which CHI has calculated an influence weight

Table 6
Comparison of research profile data with UGC rankings for physics

	Number of papers	Average influence	Influence score	Influence[a]	UGC ranking[b]
University 1	552.4	22.9	3.1	11 739	4
2	469.2	23.3	3.5	10 251	4
3	296.8	20.1	0.6	5 697	4
4	144.7	18.7	−0.1	2 679	4
5	165.2	15.5	−1.5	2 499	3
6	121.3	21.4	0.9	2 435	2
7	120.3	20.7	0.6	2 254	3
8	105.3	21.6	0.9	2 086	3
9	100.4	20.4	0.5	1 993	3
10	99.2	17.2	−0.6	1 664	2
11	95.2	17.2	−0.6	1 637	2
12	65.0	24.5	1.5	1 519	1
13	44.3	34.1	3.4	1 508	2
14	94.8	16.7	−0.7	1 466	2
15	72.6	21.0	0.5	1 426	2
16	98.2	14.4	−1.5	1 342	3
17	75.2	17.0	−0.6	1 278	N/A
18	62.4	21.6	0.6	1 125	2
19	76.1	14.8	−1.2	1 088	2
20	81.1	15.2	−1.1	1 081	1
21	69.7	15.7	−0.9	982	3
22	62.4	15.5	−0.9	947	1
23	66.1	17.3	−0.4	943	3
24	52.9	19.8	0.2	908	1
25	44.8	21.2	0.5	860	1
26	39.1	23.1	0.8	857	2
27	68.3	15.2	−1.0	842	3
28	77.3	11.8	−2.1	837	1
29	60.7	14.0	−1.3	829	N/A
30	42.1	17.9	−0.2	718	3
31	49.5	17.5	−0.3	718	2
32	36.4	19.4	0.1	706	2
33	35.2	19.4	0.1	673	3
34	53.5	14.1	−1.1	670	1
35	51.3	14.9	−0.9	650	2
36	30.3	21.1	0.4	637	1
37	38.8	17.3	−0.3	591	N/A
38	41.7	13.5	−1.2	522	1
39	30.5	17.2	−0.3	521	3
40	30.8	16.4	−0.5	505	1
41	37.7	13.4	−1.1	496	2
42	30.8	16.4	−0.5	494	1
43	28.0	16.7	−0.4	447	1
44	21.7	20.7	0.3	445	1
45	31.5	13.5	−1.0	412	1

(Table 6. cont)

	Number of papers	Average influence	Influence score	Influence[a]	UGC ranking[b]
46	54.2	8.7	−2.6	408	2
47	11.0	28.4	1.0	312	1
48	28.9	10.5	−1.5	288	1
49	20.4	14.2	−0.7	287	1
50	19.0	16.9	−0.3	270	1
51	17.3	15.8	−0.4	262	N/A
52	13.1	16.4	−0.3	215	1
53	19.4	10.4	−1.0	113	1
54	9.6	9.4	−1.0	90	N/A

[a]Universities ranked in terms of total influence.
[b]UGC rankings: 4 = outstanding
3 = above average
2 = average
1 = below average
N/A = not available.

some journals has been criticized by researchers consulted during the process of validation. Physicists, for example, pointed out that *Physics Letters B* and *Annual Review of Nuclear Science* have not been classified as 'nuclear and particle physics' – even though all the papers they contain are on this subject – but as 'general physics'. This suggests that the journal classification needs to be improved and fully validated. Second, in certain cases the influence weights of journals do not seem to accord with researchers' perceptions as to which are the more significant journals. In particular, review journals are regarded as being given undue importance. Again, more validation is required.

A third problem is that the influence indicators are significant only when considering institutions producing more than 30–40 papers in a given research area, i.e. they are generally only useful at the field rather than subfield level. Yet there can be problems in interpreting data when subfields are combined. To take a hypothetical example; suppose two universities, X and Y, each record an 'influence per paper' figure identical to the average for all UK universities in every physics subfield. However, University X publishes very heavily in nuclear and particle physics where the influence per paper for many journals is quite high (20–30), while University Y concentrates more on applied physics where most journals have an influence per paper of between 5 and 15. As a result, for physics as a whole, University X receives a large positive influence score, while University Y obtains a negative score, even though in both cases they record an influence score of zero in all physics subfields. More generally, because applied research journals tend to have lower values for 'influence per paper', the influence procedure is likely to be biassed against institutions carrying out research of a less basic nature.

Fourth, the breakdown into fields (and subfields) is, as we have seen, based on the journals in which authors publish rather than the departments in which they work. Consequently, the results for 'chemistry', for example, do not necessarily relate entirely to chemistry departments. A few of the papers in chemistry journals may have been written by, say, members of a physics department, while some chemists may publish in journals classified by CHI as 'chemical engineering'. As such, the institutional profile data are less useful for policy purposes than if the classification solely reflected departmental affiliation.[49]

Finally, if research profile data were to be adopted for routine use in policy-making, this would almost certainly affect scientists' choices as to the journals in which they elected to publish their research results. In that some institutions might pursue a strategy of publishing whenever possible in journals with the highest influence, while others might, for reasons of loyalty perhaps, refuse to alter their publication habits, this raises the question of how, if at all, one could allow for such 'manipulation' of bibliometric profile data.[50]

Nevertheless, despite all these problems and the evident need for further development work, we have seen how a combination of publication and influence profile data can yield policy-relevant information. As mentioned earlier, there is currently considerable debate on the future of British universities, with some[51] advocating that research within each field should be concentrated in perhaps 20 universities instead of spreading resources more thinly over 40–50 as at present. The results contained in Table 5 show that the bottom 50% of UK universities (i.e. the two dozen or so with the smallest output) account for only 20–25% of publications and a slightly lower proportion of total influence. However, before deciding whether to cease supporting research at such universities, several questions need to be addressed. First, what proportion of the research resources do those bottom 50% of universities consume? (If they receive only 25% of the research funds in a given field, then it would be unrealistic to expect them to produce more than 25% of the publications.) Second, if those resources were to be removed and instead concentrated on the strongest research departments, would the gains outweigh the losses? (Would there be a net addition to the number of papers published and to the overall influence?) Third, what effect would removing research from half the university departments in a particular field have on the quality of the teaching they were able to offer students? Unfortunately, all these questions were beyond the scope of the present study, but they are ones upon which future science urgently needs to cast further policy research in the UK-light.

*

The authors are grateful to Beatrice *Borer*, Lizzie *Davenport*, John *Hassett* and Francesca *McGrath* who provided research assistance at Strathclyde University, to Paul *Hackney* who wrote the software for analyzing the profile data, to Valerie *Martin* who carried out the citation analysis involved in the validation exercise, and to David *Hitchin* for statistical advice. Critical comments from Professor Richard *Cormack* were also much appreciated. John *Irvine* and Ben R. *Martin* would like to express thanks to the Economic and Social Research Council for financial support of the Programme on Science Policy and Research Evaluation.

Notes and references

1. The function of the UGC is to advise the Secretary of State for Education and Science on the distribution of government funds to the 50 or so British universities.
2. See, for example, subsequent reports and correspondence in the *Times Higher Education Supplement.*
3. In 1981, certain universities had their central government support cut by 30% or more. Subsequent years saw little or no improvement, and in 1986 universities were asked to prepare plans for the next five years based on the assumption of a reduction in public funding of 5–10% in real terms.
4. See, for example, J. IRVINE, B. R. MARTIN, *Foresight in Science: Picking the Winners,* Frances Pinter (Publishers), London and New York, 1984.
5. 'Cost centres' are in some cases larger than traditional university departments – for example, 'biological sciences' includes such departments as botany, zoology and genetics. There has been criticism of the focus by the UGC on evaluating the research output of 'cost centres' rather than departments or indeed actual research groups. It is a matter for further research whether the cost centre should be treated as the primary unit of scientific production, as is the related assumption that optimizing on this variable will also maximize the productive capacity of the university system as a whole.
6. See A. ANDERSON, Research gradings stir emotions, *Nature,* 332 (24 July 1986) 229.
7. See, for example, the list of results reported in N. CREQUER, The strengths and weaknesses, *Times Higher Education Supplement,* (30 May 1985) 4.
8. See, for example, ANDERSON, *op. cit.,* note 6.
9. See R. GILLETT, Serious anomalies in the UGC comparative evaluation of the research performance of psychology departments, *Bulletin of the British Psychological Society,* 40 (1987) 42.
10. For example, C. H. LLOYD, The research productivity of UK dental schools in the years 1980–1985, *Medical Scientific Research,* 15 (1987) 349.
11. GILLETT, *op. cit.,* note 9., p. 42. See also R. GILLETT, M. AITKENHEAD, Rank injustice in academic research, *Nature,* 327 (June 4, 1987) 381.
12. See the papers in this volume of *Scientometrics* by J. KING, and J. MCGINNETY.
13. Private communication from an official of the Economic and Social Research Council. A study exploring the potential utility of bibliometric indicators for social science research was commissioned in 1987 from PREST at the University of Manchester.
14. Additional indicators (e.g. patents) sometimes need to be employed for evaluating certain engineering and applied science specialties – see, for example, J. IRVINE, B. R. MARTIN, *Assessing the Impact of SERC Support for Engineering Research,* (mimeo), Science and Engineering Research Council, Swindon, 1985; and J. IRVINE, *Evaluating Applied Research: Lessons from Japan,* Frances Pinter, London, 1988. Bibliometric indicators should only be used with caution in these areas of research.

15. See, for example, B. R. MARTIN, J. IRVINE, Assessing basic research: some partial indicators of scientific progress in radio astronomy, *Research Policy*, 12 (1983) 61; H. F. MOED, W. J. M. BURGER, J. G. FRANKFORT, A. F. J. VAN RAAN, The use of bibliometric data for the measurement of university research performance, *Research Policy*, 14 (1985) 131; and F. NARIN, *Evaluative Bibliometrics*, National Science Foundation, Monograph 456, Washington DC, 1976.
16. G. PINSKI, F. NARIN, Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics, *Information Processing and Management*, 12 (1976) 297; and F. NARIN, Measuring the research performance of higher education institutions using bibliometric techniques, paper presented at Workshop on Science and Technology Measures in the Higher Education Sector, Paris (10–13 June 1985).
17. These are estimated to contain at least 75 per cent of all science of international relevance. However, there is an Anglo-Saxon, and, more particularly, a US bias in the *SCI*. The classification of journals was undertaken within CHI according to the main subfield of papers published in each journal. For very general areas of research, subfield categories such as 'general physics' are employed. In the case of the 100 or so journals regarded by CHI as multi-disciplinary, each journal has been classified according to the estimated subfield breakdown of the papers it contains. Details of the procedure employed are outlined in F. NARIN, M. P. CARPENTER, *Bibliometric Profiles of UK Universities and Research Institutions*, report by CHI Research to the Advisory Board for the Research Councils, London, 1987 (restricted).
18. A ninth field, psychology, was excluded from the study reported here because the transfer over time of most journals from the *SCI* to the *Social Science Citation Index* means that the coverage of the subject in the *SCI* is now unsatisfactory.
19. For details, see NARIN and CARPENTER, *op. cit.*, note 17.
20. Full mathematical details can be found in *ibid.*, pp. 74–80. An alternative approach is to solve the relevant simultaneous equations directly. However, such an approach involves more operations and is likely to be less accurate.
21. These influence weights were then applied to papers published in 1983 and 1984 by UK academic and related institutions, i.e. it was assumed that the journal influence weights did not vary appreciably in the intervening 1–2 years. This procedure can potentially lead to distortions in the case of work published in newer journals, especially those covering developing areas of research with a more applied orientation.
22. The role of the Advisory Board for the Research Councils is to advise the UK Secretary of State for Education and Science on the allocation of the 'Science Vote' among the five Research Councils.
23. More precisely, the coverage is limited to publications listed in the 1983 and 1984 *SCI* tapes. The former contains some papers dated 1982 which did not actually appear until 1983. Similarly, the 1984 tapes will have omitted publications dated 1984 but which did not appear until 1985. Since the 'slippage' is fairly constant from year to year, the two effects (of gaining some papers from the previous year and losing some to the next) are assumed approximately to cancel out.
24. P. HEALEY, H. ROTHMAN, P. K. HOCH, An experiment in science mapping for research planning, *Research Policy*, 15 (1986) 233; D. CROUCH, J. IRVINE, B. R. MARTIN, Bibliometric analysis for science policy: an evaluation of the United Kingdom's research performance in ocean currents and protein crystallography, *Scientometrics*, 9 (1986) 239; and the paper by PHILLIPS and TURNEY in this volume of *Scientometrics*.
25. See ROYAL SOCIETY, *Evaluation of National Performance in Basic Research*, Advisory Board for the Research Councils, Science Policy Studies No. 1, London, 1986; D. C. SMITH, P. M. D. COLLINS, D. M. HICKS, S. WYATT, National performance in basic research, *Nature*,

323 (23 October 1986) 681; and P. M. D. COLLINS, M. HART, D. M. HICKS, UK performance in basic solid state physics, *Physics in Technology*, 18 (1987) 72.

26. See D. HICKS, Beyond serendipity: factors affecting performance in condensed matter physics, paper presented at the 4S Annual Meeting, Worcester, MA (November 19–22, 1987).
27. Amongst the institutions excluded were colleges of further education, industrial research associations, military research laboratories, learned societies, industrial companies, and research institutes sponsored primarily by companies or charities.
28. For example, the *Commonwealth Universities Yearbook* and the various British Library directories of *Current Research in Britain.*
29. For further details, see F. GIBB, J. A. PARRISH, *ABRC Bibliometric Profiles Project*, (mimeo), Department of Information Science, University of Strathclyde, Glasgow, 1986.
30. There were in fact several stages of iteration in compiling this list involving both Strathclyde and CHI – see *ibid.*
31. For example, a paper produced by two researchers, one at at British and the other at a US university, would be credited 50% to both.
32. Irvine was at the Technical Change Centre for the period November 1986 to July 1987.
33. For each of the 499 institutions, CHI provided 8 bibliometric indicators, broken down into 8 fields and 99 subfields.
34. P. HACKNEY, *Profile Information Presentation System (PIPS)*, Science Policy Research Unit, University of Sussex, Brighton, 1987.
35. The annual editions of *SCI* were used for 1984 and 1985, and the six bimonthly versions for 1986.
36. I.e. the number of papers published in journals for which CHI had calculated influence data, multiplied by the average influence per paper for each journal, summed over all journals.
37. The results were better for the five biology groups alone with correlations of approximately 0.7, while for the five physics groups the correlation between influence per paper and citations per paper was 0.4, and that between total influence and citations was close to zero. As the above results are for product-moment correlations, it could be argued that, because the sample of universities was chosen to cover as wide a range of 'influence per paper' as possible, Kendall's rank correlation coefficient is more appropriate. However since the measurements involved are on an internal scale, and since there is no indication that the relationship between the variables is non-linear, the conditions for using the product-moment correlation are met (although the choice of universities will affect the *significance* of the correlations).
38. Even with very large samples of papers, there are at least two reasons why one would not expect perfect correlation. In the influence weight procedure, papers are fractionated, and citations are weighted according to the influence of the citing journal. Neither of these technical adjustments was undertaken in the manual citation analysis (except in the case of a number of experimental high-energy physics papers from one university where the number of collaborating institutions varied between 7 and 19, and where anomalous citation results would have been obtained if those papers had not been fractionated). For details, see B. R. MARTIN, J. IRVINE, *Bibliometric Profiles of UK Universities and Research Institutions*, a report by the Science Policy Research Unit to the Advisory Board for the Research Councils, London, 1987.
39. Compared with other UK organizations, University A is above average in terms of the bibliometric indicators shown here.
40. The number of citations to such journals was too small to be treated statistically.
41. See, for example, NARIN and CARPENTER, *op. cit.*, note 17, pp. 12–14.
42. It should be noted that the influence data for clinical medicine and mathematics for University A are probably not significant.
43. The biology category includes departments of botany and zoology.
44. As noted ea ier, caution is needed in interpreting the influence data for institutions publishing less than 30–40 papers in a given field.

45. Generally, however, it is possible to construct equivalents of the UGC cost centres by suitable aggregation of the CHI subfields. Since the process is rather complicated, such an analysis is not undertaken here, but could usefully be pursued in future research.
46. All these correlations are appreciably lower than those recorded for equivalent data in US universities – see R. ANDERSON, F. NARIN, P. MCALLISTER, Publication ratings vs peer rankings of universities, *Journal of the American Society for Information Science* 29 (1978) 91. One possible interpretation is that the UGC rankings are less reliable than the Roose-Anderson ratings of US universities.
48. A more detailed review of the methods and results is being published elsewhere by ABRC.
49. This is probably even more of a problem for data disaggregated to the subfield level because of problems with the classification of journals at that level – see the example discussed above of nuclear and particle physics. Physicists who were shown profile data for physics subfields within one university felt that they bore little relationship to their perceptions of their department's strengths and weaknesses.
50. Some might argue that such a change would not be wholly negative since the journals with higher influence tend to be regarded more highly in the scientific community.
51. For example, SIR GEOFFREY WILKINSON (see N. HALL, Nobel chemist calls for drastic closures, *Times Higher Education Supplement* (31 October 1986), p. 1); CHRISTOPHER BALL (See O. SURRIDGE, Ball calls for research elite, *Times Higher Education Supplement* (6 March 1987) p. 3); and the UGC Earth Sciences Review Committee (see UNIVERSITY GRANTS COMMITTEE, *Strengthening University Earth Sciences: Report of the Earth Sciences Review*, UGC, London, 1987).