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

Scientometrics,
Vol. 52, No. 1 (2001) 29–43

The quantity and quality of female researchers: A bibliometric study of Iceland

GRANT LEWISON

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

The output of female researchers in Iceland, relative to that of males, can be investigated because typically their "surnames" end in "dottir" whereas the names of males end in "son". Over the 21 years from 1980 to 2000, there has been a rise in female: male output from 8% to about 30%. It is higher in the life sciences (biomedical research, biology and clinical medicine) but lower where there is also foreign co-authorship, suggesting that females are less able to make overseas contacts through travel. There appears to be no difference in the quality of female and male research output, as measured either by journal impact categories or by citations.

Introduction

National science policy agenda nowadays usually include measures to promote the careers of female scientists. Thus initiatives are currently being taken in many countries: in Denmark,[1] in the European Union,[2,3] in Finland,[4] in France,[5] in Germany,[6,7] in Italy,[8] in the Netherlands,[9] in Sweden,[10] in the UK[11] and in the USA.[12] In many industrial countries females now have parity with males at the undergraduate level and even at the post-graduate level in the life sciences.[13,14] However their numbers begin to fall away at the post-doctoral level and relatively few have gained the status of full professor or academician.[15–17]

Clearly there is a need to monitor their achievements and to ascertain whether they are being treated fairly relative to males. In Sweden, for example, there was evidence of systematic discrimination in the award of Medical Research Council grants.[18] In the UK, there appeared to be equality of treatment in the award of grants[19] but far fewer females applied for grants than their relative numbers in appropriate positions would have suggested.[20]

The monitoring process is most readily conducted at or by grant-giving organisations, which will usually routinely record applicants' sex. Even if it is not specifically given, *ex post* examination can often determine the sex of an applicant from

0138–9130/2001/US $ 15.00

their first, or given, name although some are ambiguous. But to gain an impression of the relative contribution of males and females to national scientific endeavour, and thus whether the equal opportunities policies are working, requires large-scale surveys. These are costly to conduct and may suffer from unintended bias. Can bibliometrics contribute evidence that might inform policy in this area? The problem is that databases normally record only the authors' initials and not their names so that they cannot be sexed. Some journals print authors' first names but not all authors follow the practice and give them out, some preferring the more impersonal use of initials. The journal *Scientometrics*, for example, seems to have followed the convention in which male authors were given initials but females their first names until 2000, when almost all authors' first names were printed. But this is not universal practice.

There is one country, however, where family names do not really exist and people are designated by the first name of their father, followed by "son" or "daughter". This country is Iceland, a small but ancient democracy with strong Scandinavian traditions of research, especially in clinical medicine and biomedicine. Many, but not all, Icelandic scientists have names ending in "-ason", "-sson" or simply "-son" for males, and "-dottir" for females. This gives the possibility to examine Icelandic scientific papers and allocate the authors to one of three groups: males, females, or don't know. The latter category would, of course, include almost all foreigners (other than those of Icelandic descent) but it seems also to include some Icelanders.

This paper describes a small study in which this technical opportunity has been exploited to discover whether it can reveal useful facts about a country's scientific output. The span of years (from 1980 to 2000) affords a wide window through which to observe the changes in society that have led to the much greater participation of women in many professions, including science. The papers are taken from the *Science Citation Index* (SCI), a multidisciplinary database of the Institute for Scientific Information (ISI), so they can tell us about the differences (if any) between fields. Because all the addresses are given, we can determine whether females are as likely as males to co-author papers internationally, which may depend on the availability of travel grants to go to conferences abroad for their initiation. And because citation data are available, both for individual papers and for the journals in which they are published, it is possible to determine the likely impact of the papers by males and females on other scientists, and so calculate whether there is any difference in "quality".

Some caveats are in order at the beginning. First, many of the papers have authors with names that are not evidently male or female. Consequently, the fractional counting scheme that is used can only deal with a portion of the authorship. Second, Iceland is a small country (population in 1998 = 271 000) and its scientific output is still small (see

Figure 1) although it has grown very rapidly in the 1990s. This means that the data often do not have the statistical power needed to draw firm conclusions that are statistically significant. Third, the SCI may not be wholly representative of the output of the country's scientists; in particular, work on fisheries may be better communicated to those who need it (Icelandic fishermen) in ways other than in international peer-reviewed journals.

Methodology

Articles, notes and reviews with "ICELAND" in the address field were downloaded from the SCI, CD-ROM version, for the years 1980-2000. (The tally of 2000 papers was incomplete because of late processing of some papers.) They were uploaded to an MS Excel spreadsheet which contained details of authors, title, source (journal, year, volume, issue and pagination) and addresses for 2779 papers.

The journal names were categorised into nine major non-overlapping fields, see Table 1; these were based on data provided by CHI Research Inc. Papers in journals where no field was assigned, and ones in multi-disciplinary journals such as *Nature* and *Science,* were categorised on the basis of their titles and sometimes through words in their address fields.

Table 1
List of major fields, and numbers of Icelandic papers in each from 1980-2000

Field	N
Clinical medicine	1108
Earth & space	477
Biomedical research	466
Biology	315
Physics	181
Chemistry	100
Mathematics	75
Engineering & tech.	50
Social sci/other	7

The papers were also categorised by research level (RL), again on the basis of a classification provided by CHI Research Inc.,[21] and shown in Table 2 with examples of journals at each level much used by Icelandic researchers. However some journals had not been so classified and were assigned a category of zero; these accounted for only 194 papers or 7% of the total.

Table 2
Categorisation of journals by research level (RL), according to CHI Research Inc.

RL	Clinical definition	Non-clin. defin.	Examples
1	Clinical observation	Applied development	*Scandinavian Journal of Infectious Diseases, IEEE Transactions on Energy Conversion*
2	Clinical mix	Applied research	*Acta Medica Scandinavica, Aquaculture*
3	Clinical investigation	Strategic research	*International Journal of Pharmaceutics, Journal of Volcanology and Geothermal Research*
4	Basic research	Basic research	*Nature, Geophysical Research Letters*

A third categorisation of journals was by their potential impact category (PIC), with four classes based on five-year mean citation scores C_{0-4}, see Table 3. These values were obtained from ISI's Mean Expected Citation Rates file, with citation data for publications in 1988, 1990, 1992 and 1994. (Icelandic papers prior to 1988 were classified using 1988 PIC categories and ones post-1994 were based on 1994 ones.)

Table 3
Categorisation of journals by potential impact category (PIC), based on five-year mean citation scores, C_{0-4} values

PIC	C_{0-4} range	Examples
1	Below 6	*Acta Obstetricia et Gynecologica Scandinavica, Journal of Volcanology and Geothermal Research*
2	From 6 to 11	*Journal of Internal Medicine, Geophysical Research Letters*
3	From 11 to 20	*International Journal of Cancer, Physical Review B Condensed Matter*
4	Above 20	*Nature, American Journal of Human Genetics, Lancet*

The addresses were analysed by counting[*] the total numbers of addresses, D, subtracting the numbers of Icelandic addresses and so determining the numbers of foreign addresses, DF. The author field was analysed by first counting the total number of authors, A, then the numbers of Icelandic females, F, with "dottir-" in the field, then the numbers of Icelandic males, M, with "sson-" or "ason-" in the field, or just "son-" where there were no foreign addresses. The numbers of names in neither category, X, was also determined for each paper by subtraction; these are called "undetermined". Fractional counts for each paper were calculated by dividing F, M and X by A.

[*] For this purpose, a special macro kindly provided by Judit *Bar-Ilan* of the Hebrew University of Jerusalem was used.

Citation counts for some of the papers were determined over a five-year period. Because citation scores vary greatly by field, attention was focussed on just the largest field (clinical medicine; also one with a fairly high F:M ratio) and on papers published between 1988 and 1996 (n = 555). A comparison was made between citation scores and PIC values for the respective journals to see how well they were correlated. However the main purpose was to develop an alternative "output measure" in order to evaluate the relative performance of male and female researchers.

The preliminary analysis showed that the female to male aggregate ratio varied with several "input variables" such as year, main field, and whether foreign co-authors were present. Although these results were often statistically significant, it was clear that the "input variables" were not completely independent. In an effort to tease out the individual determinants of paper quality, and in particular the role of female scientists' presence in the research team, an analysis was conducted using the SPSS program and the input variables shown in Table 4. The output or dependent variable was PIC value or, for clinical medicine papers from 1988 to 1996, the citation score.

Table 4
Input variables used for the SPSS analysis

Variable	Parameters used
Authorship	A (1, 2, 3, 4, 5, 6, 7+)
Composition	M/A, F/A, X/A
Field	Eight fields (1 or 0)
Year	Year - 1979
Inter-lab cooperation	D (1, 2, 3+)
Foreign institutions	DF (0, 1, 2+)
Research level	RL, RL^2, RL^3

It was not expected that RL would have a linear effect on PIC or citation scores; the presence of both quadratic and cubic terms allowed for an arbitrary relationship between them. The effect of publication year was allowed for as it has been observed previously[22] that there is a "drift" with time towards publications in higher impact journals. This is caused in part by market forces, which cause high impact journals to expand and ones of low impact to contract or cease publication.

Results

Figure 1 shows the variation in numbers of papers per year, with a division between purely Icelandic papers and ones with a foreign co-author. (The slight reduction in output in the last year is a statistical artefact because not all the 2000 papers have been covered.) There has clearly been a big rise in output, with domestic papers tripling in number but foreign co-authored ones multiplying by about six over the period. They have gone from about 40% of the total to about 65%: this rise is typical for many small countries[23,24] and shows the increasing internationalisation of science.

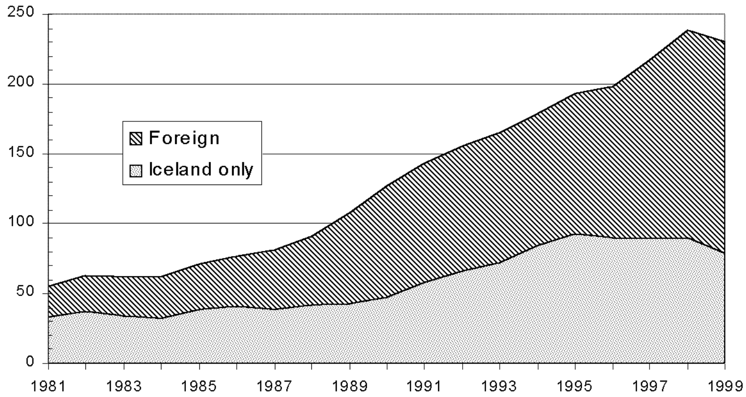

Figure 1. Icelandic scientific output in SCI, 1980-2000; three-year running means

Figure 2 shows the distribution of Icelandic papers by major field over the period, with the ones showing the greatest expansion at the bottom of the graph (chemistry and biology) and the one with the greatest relative contraction (earth and space) at the top. It is clear that the life sciences occupy a prominent place in Icelandic science and make up more than two thirds of total output. It is worth noting, parenthetically, that one of the consequences is a very high standard of public health in Iceland, with extremely low infant mortality and a high life expectancy for both males and females.

Figure 3 shows the change in the ratio of numbers of female research authorships of papers to male plus female authorships. The ratio has risen from about 7% in the 1980s to about 20% in the late 1990s. This is a dramatic increase but it appears to be levelling off in the last few years.

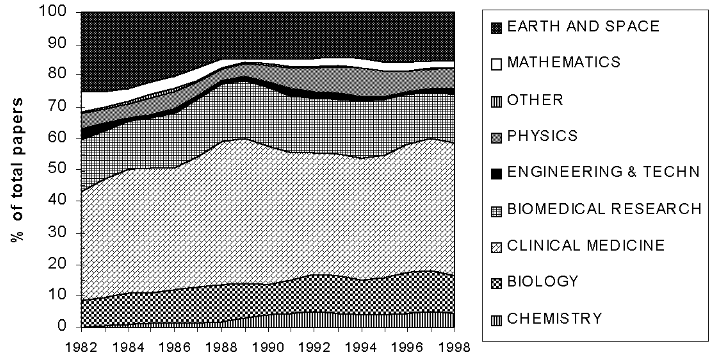

Figure 2. Distribution of Icelandic papers in SCI between main fields, 1980-2000; five-year running means

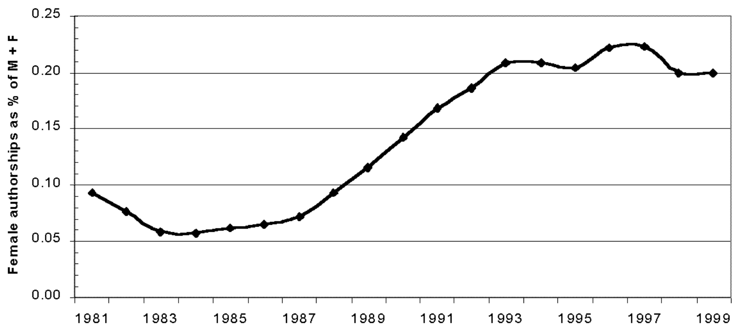

Figure 3. Ratio of female to male + female authorships of Icelandic papers in SCI, 1980-2000; three-year running means

The numbers of female researchers vary with the main scientific field, as Table 5 shows.

There is also a tendency for females to participate less in papers with foreign co-authorship. This is shown in Table 6. This table shows that the proportion of undetermined names rises sharply as the number of foreign addresses rises, as would be expected.

Table 5
Main scientific fields of Icelandic research, 1980-2000, and relative contributions of males (M), females (F) and authors with undetermined names (X), fractional counts

Field	N	M	F	X	F/(M+F)
Biomedical research	466	188.4	59.6	218.0	24.0%
Clinical medicine	1108	531.6	130.8	445.6	19.7%
Biology	315	183.3	38.3	93.4	17.3%
Chemistry	100	46.0	8.7	45.3	15.9%
Earth and space	477	220.4	23.2	233.4	9.5%
Engineering & technology	50	27.7	2.5	19.8	8.3%
Physics	181	87.4	2.4	91.2	2.7%
Mathematics	75	34.6	0.0	40.4	0.0%

Table 6
Icelandic papers analysed according to numbers of foreign addresses (DF) and their totals of male (M), female (F) and undetermined (X) fractional authorships, 1980-2000

DF	N	M	F	X	F/(M+F)	X/N
0	1209	840	199	170	19.2%	14%
1	690	283	40	367	12.4%	53%
2+	880	201	27	652	11.8%	74%

The next comparison was of the numbers of male and female authorships on papers at different research levels. The results are shown in Table 7, and suggest that females have a strong preference for work classed as "clinical investigation" or "strategic research".

Table 7
Icelandic papers analysed by research level (RL; 1 = applied, 4 = basic) and their totals of male (M), female (F) and undetermined (X) fractional authorships, 1980-2000

RL	N	M	F	X	F/(M+F)	X/N
1	194	97	17	80	15%	41%
2	726	390	73	263	16%	36%
3	692	289	89	314	24%	45%
4	973	459	65	449	12%	46%

Perhaps surprisingly, the proportion of undetermined names is relatively invariant with RL although it is slightly higher (45%) for papers at RL = 3 or 4 than for ones at RL = 1 or 2 (37%).

The results so far have shown that females participate more in Icelandic research:

- in recent years;
- in the life sciences;
- in domestic projects;
- in clinical investigation or strategic research work.

In order to see if these factors interact, and if there are other important factors, an analysis of the papers was conducted using the SPSS program with F/A as the dependent variable, and the independent variables being:

A, the number of authors;

DF, foreign addresses;

YR, Y2, Y3, the year of publication – 1979, its square and its cube, to allow for a non-uniform relationship with time:

BI, BM, CH, CM, ES, ET and PH, the main fields;

RL, RL2, RL3, functions of the research level to allow an arbitrary relationship.

The results are shown in Table 8, and the implied relationship with year in Figure 4.

Table 8
SPSS analysis of fractional female participation in a research paper, F/A, as a function of other independent variables for Icelandic research papers, 1980-2000

Parameter	Coefficient	Significance	Parameter	Coefficient	Significance
Authors, A	0.014	< 0.01%	Biology	0.093	0.01%
Addresses, D	0.001	n.s.	Biomed res	0.105	< 0.01%
Foreign addr, DF	-0.091	< 0.01%	Chemistry	0.047	n.s.
RL = 1 clinical	-0.048		Clin med	0.088	0.02%
RL = 2	-0.027		Earth & space	0.044	5%
RL = 3	0.005		Engr & tech	0.049	n.s.
RL = 4 basic	-0.026		Physics	0.010	n.s.

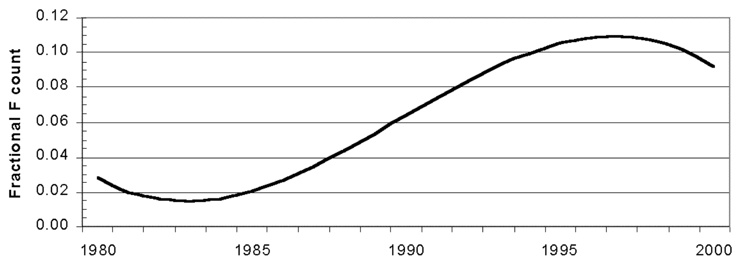

Figure 4. Variation of female participation in Icelandic research with year, assuming a cubic relationship and analysis by SPSS

Figure 4 bears a striking resemblance to Figure 3, with an initial decline in the early 1980s, a steady rise to a maximum around 1996 and a small decline thereafter. Table 8 confirms the trends seen when the variables were investigated singly, with maximum participation of females in biomedicine, followed by biology and clinical medicine, and less participation if there are foreign collaborators. But it also shows that females are more likely to participate if the research teams are larger. This is highly plausible as they only make up a small fraction (less than one tenth) of the total authorship. Since authorship teams have increased steadily in size with time (2.6 in 1980-84; 3.4 in 1985-89; 4.5 in 1990-94; 5.5 in 1995-99), this is another reason why female participation in research has increased.

The next matter to be investigated was the perceived relative "quality" of their work, as determined by PIC values and by citations. Overall, the distribution of papers between PIC categories is extremely similar for males and females as Figure 5 shows.

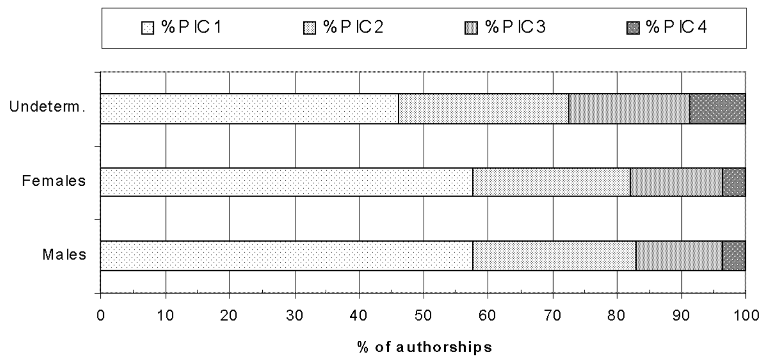

Figure 5. Distribution of Icelandic research papers in SCI by potential impact category (PIC; 1 = low, 4 = high) for male, female and undetermined authorships (fractional counts), 1980-2000

However it is clear that the undetermined authorships, mainly foreigners, are producing work of higher potential impact. The difference is highly statistically significant (χ^2 = 59.5, 3 d/f). This does not mean, however, that Icelandic scientists, both male and female, are producing low impact work relative to that of the foreigners.

It is known[24,25] that internationally co-authored papers tend to have higher impact than domestic ones and nearly all (86% on a fractional count basis) of the undetermined authorships are in papers with foreign addresses. On the other hand, only 36% of the male and female authorships are in such papers.

The data (minus the 7 papers in social sciences /other fields) were analysed using the SPSS program and the independent variables listed in Table 4. The results are shown in Table 9 for PIC as the dependent variable. Although only 21% of the variation could be explained by the given input variables, most of them had coefficients that were statistically significant.

Table 9
SPSS analysis of potential impact category (PIC) as a function of various input variables for Icelandic papers in SCI, 1980-2000

Variable	Coeff.	p	Variable	Coeff.	p
Authors, A	0.084	< 0.01%	Biology	0.217	4%
Addresses, D	0.004	n.s.	Biomed res	0.684	< 0.01%
For'n addr, DF	0.046	n.s.	Chemistry	0.319	1%
Year	0.009	0.3%	Clinical medicine	0.802	< 0.01%
RL = 1 applied	0.443		Earth & space	0.719	< 0.01%
RL = 2	0.498		Engr & tech	0.319	4%
RL = 3	0.606		Mathematics	n.a.	
RL = 4 basic	1.109		Physics	0.431	0.01%

The coefficients for the parameters M/A and F/A were –0.12 and –0.19 respectively, but neither was significant at the 5% level. This means that the difference in PIC for papers with male and female authorships (as opposed to undetermined ones) is not large enough to be statistically significant although the mean PIC for females is a little lower than for males when the other input factors have been accounted for.

The distributions of citations to the 555 papers in clinical medicine published from 1988 to 1996 are shown on a log-log graph in Figure 6, where the abscissa (x-axis) represents the centile of each of the groups. The dotted line represents the undetermined group (N = 216.1 papers), the dashed line represents the female authorships (N = 74.3 papers) and the solid line the male authorships (N = 264.6 papers). The papers are fractionated according to the contribution that each type of author has made to the total. The numbers of citations obtained by papers at the 1st, 2nd, 3rd, … 40th, 50th, 60th, 70th centiles are determined from a ranking of all the papers with undetermined, female and male authorships.

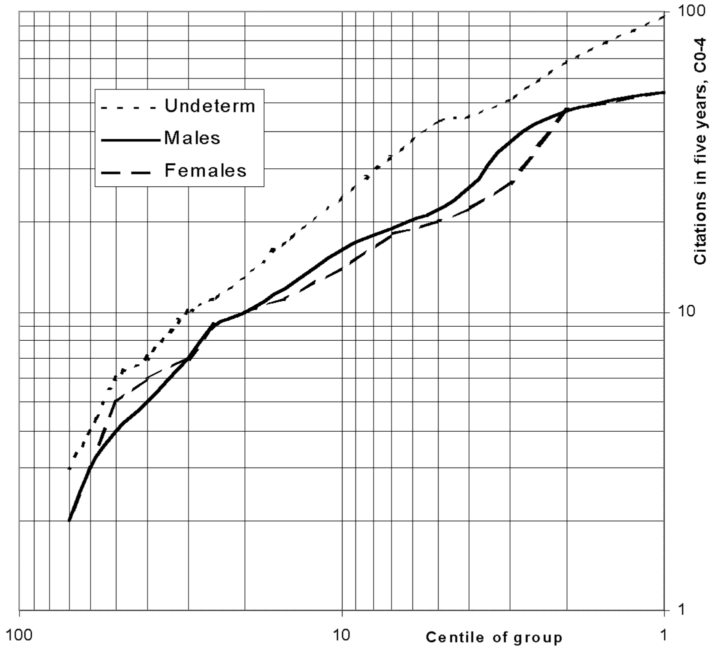

Figure 6. Distribution of five-year citations to Icelandic clinical medicine papers, 1988-96, divided according to fractional contribution of male, female and undetermined authorships

The result is rather similar to that of Figure 5, with the undetermined authorships having the most highly cited papers (top decile, or tenth centile, has 24 cites compared with 16 for male authorships and 14 for female ones). There is little difference in the citations to male-authored and to female-authored papers.

The 555 papers with citations were analysed using the SPSS program to see what factors might influence citation scores, in particular whether there was any difference between male and female impact. For this analysis, PIC was treated as an independent variable. The results are shown in Table 10. In this table the only significant variables were year and PIC, the latter being very highly significant as would be expected. The effect of research level was small for papers at RL 1, 2 and 3 but appeared to be rather negative for RL = 4. However since there were only 15 papers in this group, with a total of 111 citations (7.4 per paper), the result is not significant.

Table 10
SPSS analysis of factors influencing five-year citation scores for 555 Icelandic papers in clinical medicine, 1988-1996

Parameter	Coefficient	Standard error	Significance
Authors, A	0.706	0.437	n.s.
Male fraction, M/A	-1.852	2.677	n.s.
Female fraction, F/A	-0.921	3.574	n.s.
Addresses, D	-0.465	0.906	n.s.
Foreign addresses, DF	0.698	0.941	n.s.
Year	0.540	0.254	3%
PIC value	6.344	0.744	< 0.01%

The relationship between citation scores for individual papers and the PIC of the journal in which they were published can be seen in Figure 7. Although publication in a high PIC journal does not, of course, guarantee that the paper will be well-cited, it certainly makes it much more likely.

There is also a positive coefficient for year, which suggests that citations to Icelandic medical papers have increased in recent years faster than would have been expected from the "drift" in journal PIC values discussed earlier. This indicates that the impact of such research has increased steadily, even when account is taken of factors such as the increasing numbers of authors and addresses.

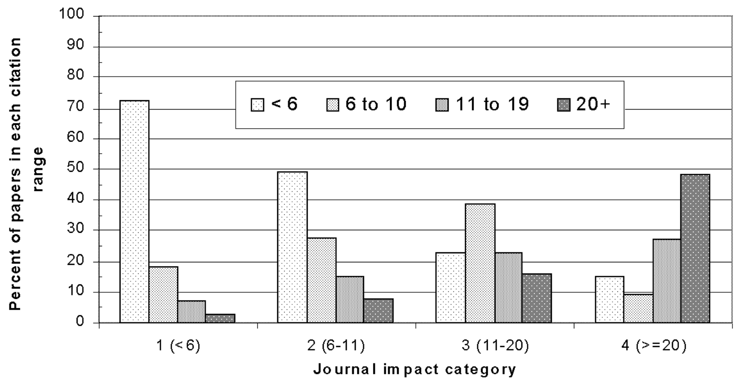

Figure 7. Five-year citation ranges for Icelandic clinical medicine papers (N = 555) published in journals of different potential impact categories

Discussion

The analysis has shown how female researchers have increased their presence in one country over the last two decades, but really only in the life sciences. There appears to be a flattening off in recent years in the percentage of papers with female authorships, suggesting that additional measures may be needed for them to achieve parity of output with men. In particular, the relative paucity of female authorships in papers with foreign addresses suggests that more could be done to help them travel abroad to conferences and for study tours. There is also little evidence of female scientific activity in physics and mathematics. The data are too few to reveal any significant differences in potential impact (journal citation scores) or actual impact (citation numbers) between papers with male and female authorship, but they do show that papers with undetermined authorship (mainly foreigners) are of higher impact as would be expected.

In this unique situation, bibliometrics has been able to provide useful information for policy purposes. It would be very difficult to extend such a study to other countries unless there were a complete list of the country's researchers available with their sexes so that comparisons could be made. An alternative would be to examine the original papers so as to try and sex the authors whose first names were printed on the paper. However this would need to be done for the whole of the country's output in particular fields as coverage of only a sample of journals might lead to unintended bias in the results. Such an exercise would only be economically feasible if combined with inspection of the papers for other reasons, such as to identify funding sources from the acknowledgements.

*

The citation analysis was carried out by Adriana *RoaCelis-Atkinson*, while visiting City University as part of her PhD training at Campinas University in Brazil. The author is a senior policy adviser of the Wellcome Trust on secondment to City University.

References

1. DE LAINE, M., Danish women fight to bring home the bacon, *Times Higher Education Supplement*, Issue 1318, 6 February, 1998, p. 11.
2. ABBOTT, A., Europe seeks greater role for women on key advisory panels, *Nature*, 399 (1999) 290.
3. ANON, Women's group given mission, *Research Europe*, Issue 71, 3 February, 2000, p. 3.
4. SEPPANEN, R., Male culture holds back women, *Times Higher Education Supplement*, Issue 1384, 14 May, 1999, p. 13.
5. SANDERS, E., Fresh ideas to help women into science, *Physics World*, April, 2000, p. 9.

6. ABBOTT, A., Increase in German science budget a boon for women and youth, *Nature*, 399 (1999) 186.
7. ANON., Germany promotes equal opportunities, *Outlook on Science Policy*, 22 January, 2000, pp. 5–6.
8. ANON., Italy adopts measures to promote women scientists, *Research Europe*, Issue 77, 4 May, 2000, p. 7.
9. ANON., Dutch to "sow and reap" in research, *Outlook on Science Policy*, 21 October, 1999, pp. 105–106.
10. ANON., Court challenge for Sweden's women professors, *Research Europe*, Issue 66, 4 November, 1999, p. 6.
11. LODER, N., British drive for more women in science, *Nature*, 398 (1999) 4.
12. COX, A. M., WILSON, C., Leaders of nine universities pledge to improve conditions for female scientists, *Chronicle of Higher Education*, XLVII, 9 February, 2001, p. 12.
13. SIDDALL, M. E., Women in biological sciences, *Science*, 283 (1999) 1116–1117.
14. HODEN, C. (Ed.), Women inch toward parity in science, *Science*, 283 (1999) 323.
15. BUHLMAN, E., Women in science in Germany, *Science*, 286 (1999) 2081.
16. ANON., 8.1% of professors are women, *Times Higher Education Supplement*, Issue 1339, 3 July, 1998, pp. 18–19.
17. HEALY, B., Women in science: from panes to ceilings, *Science*, 255, (1992) 1333.
18. WENNERAS, C., WOLD, A., Nepotism and sexism in peer review, *Nature*, 387 (1997) 341–343.
19. GRANT, J., BURDEN, S., BREEN, G., No evidence of sexism in peer review, *Nature*, 390 (1997) 438.
20. BLAKE, M., LAVALLE, I., *Who applies for research funding?* The Wellcome Trust, London: ISBN 1 841290 29 7, 2000.
21. NARIN, F., PINSKI, G., GEE, H. H., Structure of the biomedical literature, *Journal of the American Society for Information Science*, 27 (1976) 25–45.
22. LEWISON, G., DEVEY, M. E., Bibliometric methods for the evaluation of arthritis research, *Rheumatology*, 38 (1999) 13–20. See Figures 2, 3.
23. INGWERSEN, P., On-line indicators of Danish biomedical publication behaviour 1988-96: international visibility, impact and co-operation in a Scandinavian and world context, *Research Evaluation*, 8 (1999) 39–45.
24. GLÄNZEL, W., Science in Scandinavia: a bibliometric approach, *Scientometrics*, 48 (2000) 121–150.
25. NARIN, F., STEVENS, K., WHITLOW, E. S., Scientific co-operation in Europe and the citation of multinationally authored papers, *Scientometrics*, 21 (1991) 313–324.

Received April 19, 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