This is a **postprint version** of: Mauleón, E.; Hillán, L.; Moreno, L.; Gómez, I.; Bordons, M. “Assessing gender balance among journal authors and editorial board members.” **Scientometrics** 95(1): 87-114, 2013. DOI: 10.1007/s11192-012-0824-4
The final publication is available at Springer: http://link.springer.com/article/10.1007/s11192-012-0824-4

The effect of gender on research staff success in life sciences in the Spanish National Research Council (CSIC)

Elba Mauleón
Centro de Información y Documentación Científica (CINDOC), CSIC, Madrid, Spain (elba@cindoc.csic.es)

María Bordons
Centro de Información y Documentación Científica (CINDOC), CSIC, Madrid, Spain (mbordons@cindoc.csic.es)

Charles Oppenheim
Department of Information Science, Loughborough University, Loughborough, UK (C.Oppenheim@lboro.ac.uk)

Abstract

Inter-gender differences in research performance of CSIC scientists in the area of Biology and Biomedicine are analysed by means of bibliometric indicators (SCI, 1996-2000). Productivity of both men and women increases with scientific category, and inter-gender differences are not found within each category. Women with intermediate levels of seniority (11-20 and 21-30 years of working life) show lower productivity than their male counterparts, factor which might contribute to the lower promotion observed for female scientists. However, women that entered CSIC in the last 10 years overpass men in productivity, so a more balanced distribution of women by scientific ranks would be expected in the future. The need to improve the normalisation of author names in publications and bibliographic databases and even to develop a “digital author identifier” to make these studies easier is pointed out.

Keywords: Citation analysis; gender studies; women scientists; women and science

Introduction

There is a growing concern in many countries about the under-representation of women in scientific careers and research. A clear gender imbalance in the sciences has been described at the European level: women make up half the student population, but hold only 15% of senior academic positions. These figures change according to fields and countries, but the same trend is observed in different geographical regions such as Europe (She figures, 2006) or the United States (Lawler 2006), while the situation is even more unbalanced in most Asiatic countries (UIS Bulletin 2006).

Within the European Union, different initiatives have been developed to analyse the presence of women in science and promote their participation. The Women and Science action plan was set out by the European Commission in 1999, and the ETAN report on women and science was produced to review the position of women in higher education, research institutes and industry at both the EU and the member state level (ETAN, 2000). This report put forward that women were underrepresented in science, promoted more slowly and received less recognition than their male counterparts. Two main ways of action were recommended. On the one hand, the need to collect precise and reliable data regarding the situation of women in science, education and technology; on the other, the need for proper identification and further elimination of barriers and inequalities that tie women to certain scientific fields and also limit women's access to the top ranks in the scientific career. Since then, different studies and publications have been carried out to see how the participation of women is evolving in science and technology (WIR, 2003, ENWISE, 2003, She Figures 2006).

To reveal the status of women in science, national sex-disaggregated statistics are needed, as well as gender sensitive indicators in order to monitor the participation of women in research (ETAN, 2000). Although most R&D indicators are input oriented, it is also possible to analyse the output of research by gender. Bibliometric indicators, based in the analysis of scientific publications, are widely used for assessing research performance of scientists assuming that scientific publications are an important output of scientific activity. As the EU suggested through the Helsinki Group, the elaboration of bibliometric indicators by gender should be promoted since these indicators could provide comparable and objective information regarding research performance from the point of view of gender.

For these reasons, studies on women's participation in scientific knowledge production have become increasingly important. This line of research is complex due to the limited amount of previous work conducted on the subject that provided an effective and standardised methodology. Moreover, difficulties are encountered in the identification of a researcher's sex through his/her name as recorded in a publication, as most publications and databases only include the initials of the first names.

Aware of these obstacles, the EU commissioned two studies to analyse women's contribution to science by means of bibliometric and patent indicators by gender (Biosoft, 2001). The methodology was to create a file with names in different languages (German, Spanish, French, English, and Swedish) in order to conduct studies by gender based on researchers' names. However, this approach involves resorting to the original publications, which frequently fail to include the authors' full name. The researchers found that two-thirds of the publications under study did not include this information, which forced them to limit the analysis to works that did (not necessarily a valid sample). This shortcoming limits the usefulness of the method used.

In some countries, it is possible to determine the sex of authors from their surnames, which contain gender information. Polish surnames, for instance, can be determined as being male or female in 60% of the cases (Webster, 2001). Similarly, the gender of most Icelandic researchers can be identified from their surnames (Lewison, 2001). However, this particular feature can be found in very few (mainly Slavic) countries.

Since gender recognition through the signature is often difficult, one alternative is to contact the authors directly; conducting surveys or interviewing them (Kyvik and Teigen, 1996; Prpic, 2002), but the success of this approach depends on the rate of response. Another option is to use researchers' curricula vitae (CV) as a source of information (Hemlin and Gustafsson, 1996). This is the best option for those institutions and countries with institutional, regional or national databases with standardised and updated CVs of scientists, but not all have this valuable source of information.

The problem described above highlights the need to develop alternative methodologies and approaches to the incorporation of the gender dimension in bibliometric studies. Focusing the analysis in specific centres or institutions, for which initial information regarding the centre's staff is available (Long, 1992; Bordons et al, 2003; Sánchez Peñas and Willet, 2006), may be an interesting alternative to cope with the problem of authors' names. Once the full name, working address and research area of the scientists are known, it is easier to search, collect and analyse their publications as covered by bibliographic databases. This is the methodology followed in this study.

Female scientists at the Spanish National Research Council (CSIC)

The Spanish National Research Council (CSIC) is the largest public research body in Spain. It is a multidisciplinary institution devoted to promote and conduct research in all fields with the aim of contributing to the advance of science as well as to the economic, social and cultural development of the country. In 2002, the Spanish Research Council (CSIC) had 2252 permanent scientists, of whom 717 (31.8%) were female. This percentage is very similar to that found in the Spanish university system (32%) (González Duarte, 2004), in the French CNRS (30%) (Crance, 2002) or in the Italian CNR (31%) (Palomba and Menniti, 2002).

CSIC is organised into eight scientific and technical areas. A detailed analysis of the presence of women in the institution shows an unequal distribution over disciplines. The lowest percentage of women is found in Physics (20%), followed by Natural Resources (25%) and Biology and Biomedicine (30%), while the highest percentage is in Food, Science and Technology (44%). Intermediate values are found in Agriculture, Chemistry and Social Sciences/Humanities (36%) (Figure 1). The different presence of women by areas is called "horizontal segregation" (concentration of women in specific fields), while differences by professional rank are known as "vertical segregation" (concentration of women in the lowest ranks within the academic ladder). Inter-gender differences in the distribution of scientists in CSIC by professional rank are observed: 38% of Tenured Scientists (lowest rank) are women vs. 28% of Research Scientists (intermediate rank) and 14% of Research Professors (highest rank), with differences by areas. The graphical representation of the percentage of men and women by professional ranks is known as the "scissors diagram", in which the percentage of women decreases as we go up in the professional rank (Figure 2).

Figure 1. Distribution of CSIC scientists by gender

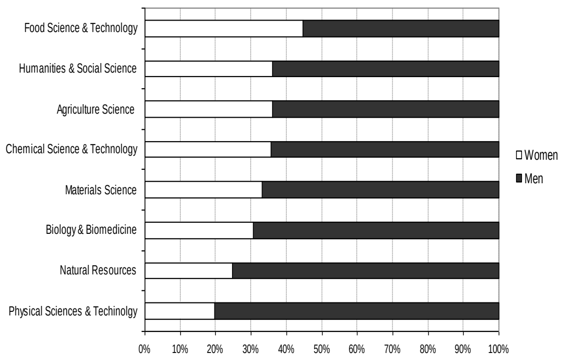
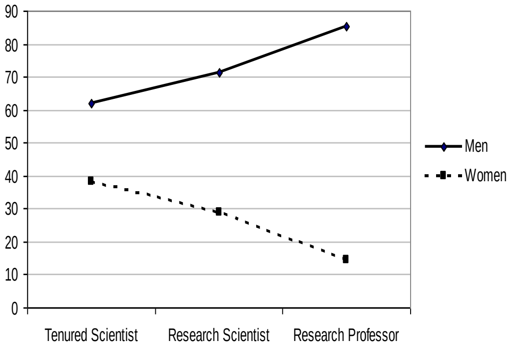


Figure 2. Percentage of male and female scientists within each professional rank at CSIC



The scissors diagram applies to many research centres and universities in different countries. The Centre National de la Recherche Scientifique (CNRS, France), for example, is a public research centre with characteristics similar to those of the CSIC: the proportion of female researchers in 2005 was 31% (Crane, 2002), but at the level DR1 (Directeur de Recherche, or Research Professor at the CSIC), women barely exceeded 10%.

Objectives of our research

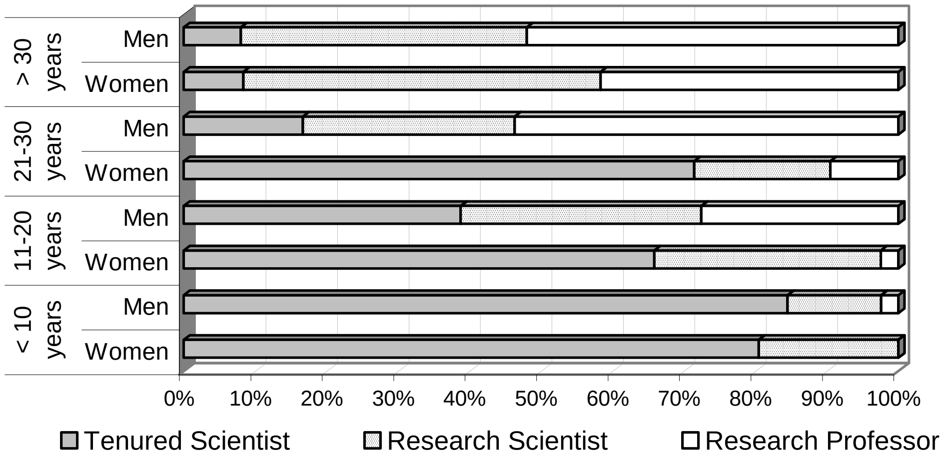
This paper describes research on Biology and Biomedicine (BIOL) at CSIC. The percentage of women in the area of Biology and Biomedicine is 30%, very close to the average at CSIC (32%). The percentage of women in Biology and Biomedicine decreases as we go up in the scientific category: 39% of Tenured Scientists, 31% of Research Scientists and only 10% of Research Professors are women ($Chi^2 = 21.77$; $p<0.001$).

A common explanation for the under representation of women in science is that they entered the scientific profession later than men. According to this hypothesis, it is a just a question of time before women arrive at the upper positions, since a natural readjustment will occur as time passes and a higher number of women enter science. Since age increases with professional rank, the present low number of women in the upper categories can be due to the low percentage of women in the age bracket of those who have arrived now at the upper positions. However, our hypothesis is that if that were the case, we would expect a similar distribution of men and women over professional ranks when controlling for age or for number of years at the institution.

The comparative distribution of men and women by scientific category when controlling for number of years at CSIC in Biology and Biomedicine research area, is shown in Figure 3. As the number of working years at CSIC increases, the percentage of scientists in the upper categories also goes up for both men and women. However, it seems that the promotion of women occurs at a slower pace than that of men.

No woman with less than 10 years at CSIC has arrived at the upper category, while 2% of men are Research Professors. Among scientists with 11-20 years at CSIC, 65% of women remain in the lowest category vs. 39% of men. Finally, 71% of women with 21-30 years at CSIC remain in the lowest category, while only 17% of men in this professional age group are still Tenured Scientists.

Figure 3. Distribution of BIOL research personnel by number of years at CSIC, sex and scientific category



At CSIC, as well as in other meritocratic systems within the academic world, scientific publications are a key element in the promotion of scientists, so bibliometric indicators by gender can shed light on possible inter-gender differences in the scientific activity of men and women, which in the long term could contribute to explain the absence of women at the highest levels of hierarchy at CSIC.

The main objective of this study was to analyse inter-gender differences in scientific activity by means of bibliometric indicators which could contribute to explain the under-representation of women in the top ranks. The following questions were addressed:

- Are there inter-gender differences in productivity? Are men more productive than women?
- Are there differences in the impact of the research of male and female scientists? Do male publications attain higher impact than female ones?

To answer these questions, the scientific publications of CSIC scientists in the area of Biology and Biomedicine during 1996-2000 were analysed by gender, scientific category and number of years at CSIC. This study forms part of an ongoing research project focused on obtaining bibliometric indicators by gender at CSIC. Studies dealing with the analysis of Chemistry, Natural Resources and Materials Science have been published elsewhere (Bordons et al 2003; Mauleón and Bordons, 2006).

Methodology

Sources of information

Two different types of information were used in this study: data on research personnel and on scientific publications from CSIC scientists in the area of Biology and Biomedicine.

a) Research personnel

For every permanent scientist at CSIC, the following data were obtained from the Human Resources Department of the institution: full name (first name and two surnames), sex, number of years at CSIC as permanent scientist, scientific category, and research centre. The scientific career at CSIC comprises three different scientific categories: Tenured Scientist, Research Scientist and Research Professor.

b) Scientific publications

There is no institutional database with the CVs and publications of CSIC scientists. Scientific publications of CSIC scientists during 1996-2000 were downloaded from the Science Citation Index database (CD-ROM version) (Thomson Scientific, Philadelphia). This database includes more than 3,600 journals selected according to quality criteria in scientific content and formal issues. Citations received by documents since their publication year to June 2005 were retrieved from the Web of Science database and added to the bibliographic records (Costas and Iribarren-Maestro, 2007).

The lack of consistency of author names is an important problem in bibliographic databases, since the scientific production of a given author can be scattered under different variant names. Normalisation of data, always necessary in bibliometric studies, is extremely important in bibliometric studies at the micro level, since higher levels of precision are needed. To overcome this problem and assure a comprehensive retrieval of scientists' publications, a detailed search strategy for publications was developed.

For each CSIC scientist, a list with all his/her different possible ways of signing in documents was built, following different studies for the construction of variant names depending of the original structure of scientists names (Ruiz-Pérez et al. 2002; Costas and Bordons, 2007). For example, publications of a scientist called "Abadia Bayona, ME" were searched under the following different author names: AbadiaBayona ME, AbadiaBayona M, Bayona MEA, Abadia ME, Abadia M, Bayona MA, AbadiaBayona E, Abadia E, Bayona MA. To deal with homonyms, addresses of authors were taken into account and potential publications of a given author were compared in relation to co-authors, publication vehicle, and research topics. Checks on the web pages of scientists and even with the scientists themselves were carried out to assure the correct identification of authors' publications.

<u>Bibliometric profile of scientists</u>

Once the authors´ names had been standardised, a bibliometric profile was built for each scientist, including different indicators to study his/her productivity and visibility:

a) Productivity

- Number of publications. As a proxy for the productivity of scientists, their number of publications in the international and prestigious Science Citation Index database (SCI, CD-ROM version) was considered.

Different methods for allocating credit to individuals have been described. The "straight count" approach allocates all credit only to the first author of a paper. In the "fractional count" approach, each author is given a credit equal to $1/a_i$ (a_i is the number of authors). The normal or "total count" gives full credit to all contributors regardless of the order of the listed authors. In this paper, the total count method was used. All types of documents were considered.

b) Visibility. The impact of documents in the international scientific community was studied through impact factor and citation-based indicators. Impact factor of journals in 2000 (JCR, 2000) and citations since the publication year to June 2005 were considered.

- Average impact factor: average impact factor of publication journals weighted according to the number of documents published in each journal was used.
- Number of citations per document.
- Percentage of non-cited documents.

- Percentage of highly cited documents: the percentage of documents included among the 10% most cited CSIC biological documents (90 percentile).

SPSS, version 12, was used for the statistical analysis of data. Kruskal Wallis and Man-Whitney tests for the comparison between means of non-parametric distributions were applied.

Results

Scientific productivity and impact by gender

A total of 336 scientists (89% of the scientists) had at least one SCI publication in the period under analysis. Only 28 men (9% of men) and 12 women (10% of women) had no publications. A total of 2925 publications were identified for the 336 scientists.

On average, each scientist published 12 documents. Average productivity of women was around 10 documents, vs. 12 documents for men, but differences by gender were not statistically significant (Table 1). Concerning impact factor of publication journals, as a proxy for the prestige of publication journals, inter-gender differences were not found.

334 scientists (99% of scientists with publications) had at least one cited publication, and 2675 publications (91% of publications) received at least one citation. No inter-gender differences were found.

On average, each scientist received 23 citations per document, and no inter-gender differences were found in the number of citations per document (Table 1).

The 10% most cited documents in the area received at least 50 citations per document. Scientists had 9% of their publications among the 10% most cited, being this value higher for men (10%) than for women (7%) ($p<0.05$).

Table 1. Productivity and visibility of BIOL scientists by gender (SCI, 1996-2000)

	Female (N=103)	Male (N=233)	Total (N=336)	Sig
No. SCI documents	9.75±7.74 (1.00-53.00)	12.64±11.32 (1.00-79.00)	11.75±10.43 (1.00-79.00)	NS
Av.IF	4.97±2.72 (0.92-16.44)	5.27±2.94 (0.69-19.68)	5.18±2.88 (0.69-19.68)	NS
No citations/document	20.25±17.09 (0.00-113.63)	24.38±23.81 (0.00-250.33)	23.12±22.03 (0.00-250.33)	NS
% highly cited documents	7.32±13.90 (0.00-86.67)	10.27±16.33 (0.00-100.00)	9.36±15.66 (0.00-100.00)	0.05
% non cited documents	7.12±15.18 (0.00-100.00)	6.87±11.98 (0.00-100.00)	6.95±13.02 (0.00-100.00)	NS

Figure 4. International productivity of BIOL scientists by gender

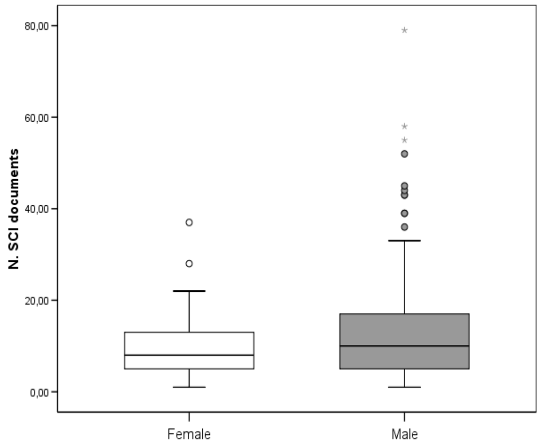


Figure 5. International visibility of BIOL scientists by gender

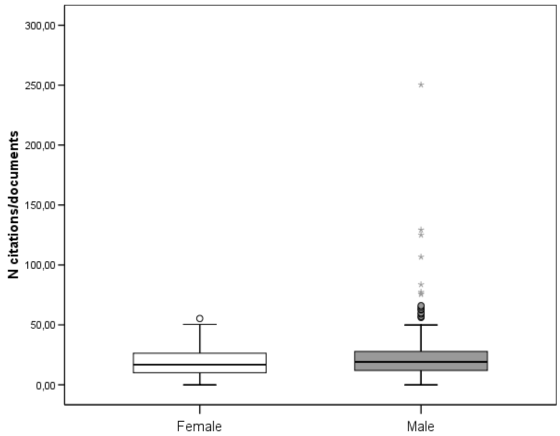


Figure 6. Percentage of highly cited documents of BIOL scientists by gender

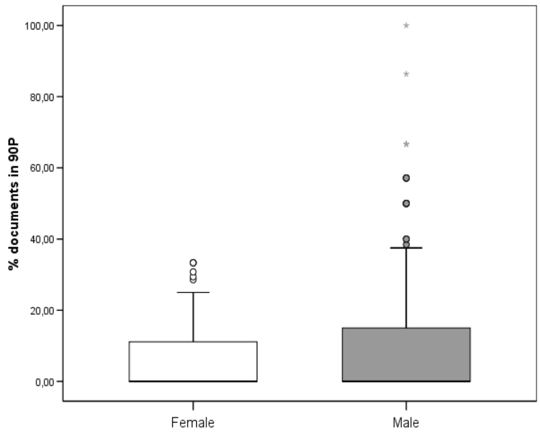
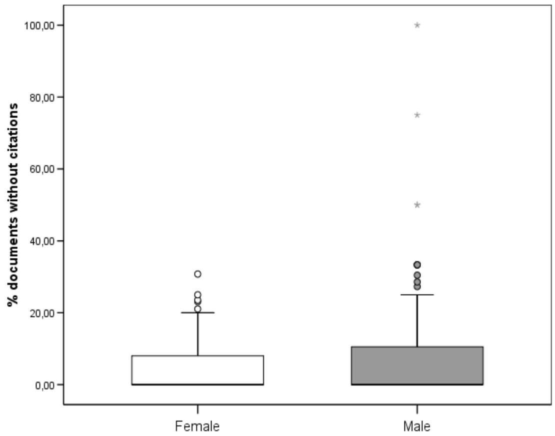


Figure 7. Percentage of non cited documents of BIOL scientists by gender



Thus, although female values were slightly lower than males ones, no significant differences were found between male and female productivity and impact. However, the maximum values of productivity, number of citations per document and percentage of highly cited documents all went to men.

<u>Scientific productivity and impact by scientific category and gender</u>

Previous studies have shown differences in productivity of scientists by scientific category: productivity tends to increase with rank. This is not surprising, since scientific publications are an important output of the research performance of scientists and promotion decisions rely heavily on them. Here, the relationship between these variables is analysed.

Does productivity increase with scientific category for both men and women? Table 2 and Figure 8 show that the productivity of both men and women increases with rank ($p<0.01$).

Are there inter-gender differences in productivity within each scientific grade? Differences at this level could suggest the existence of gender discrimination in promotion. If the same level of achievement is required for both men and women to be promoted, no differences between genders would be expected. Our results show that within each scientific grade, inter-gender differences in productivity are not statistically significant.

Whilst productivity increases with rank (Figure 8), this is not the case of average impact factor and citations, which do not show a clear relationship with rank (Table 2, Figure 9). Men got a higher number of citations per document than women in the Tenured Scientist category ($p<0.05$), while no differences by gender were found in the other two categories.

Whatever their rank, 9-10% of scientists publications were highly cited documents; and inter-gender differences were not found. Neither were differences found in the percentage of non-cited documents published (6-8%).

Figure 8. Productivity of BIOL scientists by scientific category and gender

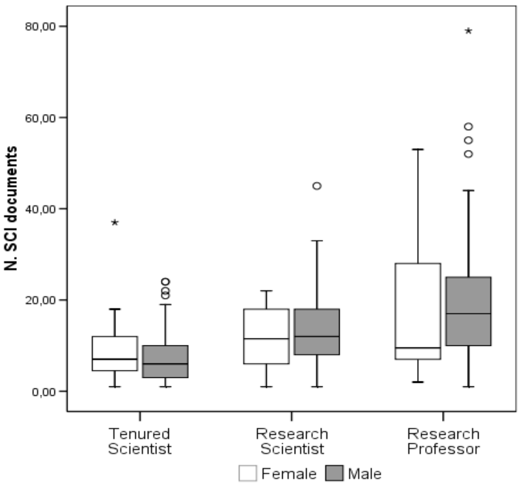


Figure 9. International visibility of BIOL scientists by scientific category and gender

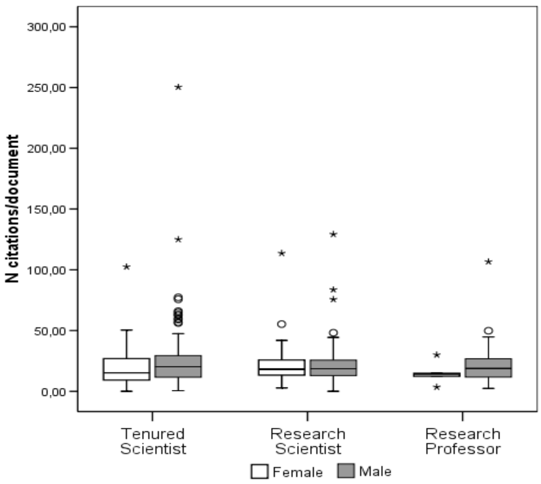
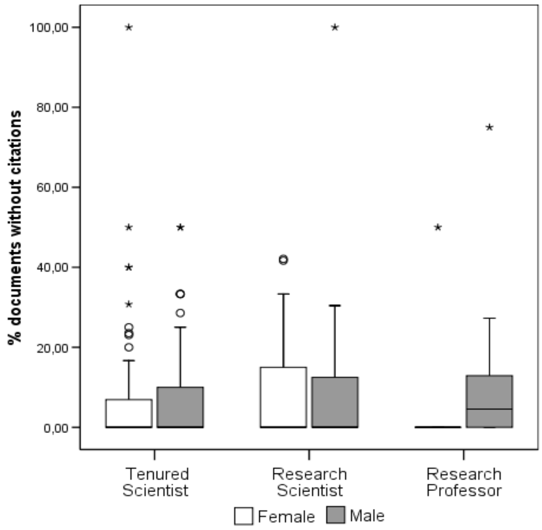


Table 2. Productivity and international visibility of BIOL scientists by scientific category and gender

	Scientific category	Female	Male	Total	SIG
No SCI documents	Tenured Scientist	8.25±6.01 (1.00-37.00) N=67	7.30±5.22 (1.00-24.00) N=105	7.67±5.54 (1.00-37.00) N=172	NS
	Research Scientist	11.40±6.51 (1.00-22.00) N=30	13.31±8.89 (1.00-45.00) N=61	12.68±8.19 (1.00-45.00) N=91	NS
	Research Professor	18.17±19.22 (2.00-53.00) N=6	20.37±15.14 (1.00-79.00) N=67	20.19±15.36 (1.00-79.00) N=73	NS
Av. IF	Tenured Scientist	4.83±2.69 (0.92-15.08) N=67	5.58±3.33 (0.69-19.68) N=105	5.28±3.11 (0.69-19.68) N=172	NS
	Research Scientist	5.26±2.95 (1.34-16.44) N=30	4.81±2.39 (1.29-11.81) N=61	4.96±2.58 (1.29-16.44) N=91	NS
	Research Professor	5.21±2.14 (2.84-7.95) N=6	5.20±2.72 (1.31-16.85) N=67	5.20±2.67 (1.31-16.85) N=73	NS
No citations/document	Tenured Scientist	19.23±15.83 (0.00-102.47) N=67	26.99±29.38 (0.50-250.33) N=105	23.97±25.22 (0.00-250.33) N=172	0.05
	Research Scientist	23.63±20.59 (2.83-113.63) N=30	22.98±20.41 (0.00-129.23) N=61	23.20±20.36 (0.00-129.23) N=91	NS
	Research Professor	14.84±8.56 (3.50-30.00) N=6	21.57±15.27 (2.33-106.68) N=67	21.01±14.91 (2.33-106.68) N=73	NS
% highly cited documents	Tenured Scientist	6.58±13.73 (0.00-86.67) N=67	11.43±19.31 (0.00-100.00) N=105	9.54±17.47 (0.00-100.00) N=172	NS
	Research Scientist	8.76±13.23 (0.00-52.63) N=30	8.62±12.08 (0.00-57.14) N=61	8.67±12.39 (0.00-57.14) N=91	NS
	Research Professor	8.33±20.41 (0.00-50.00) N=6	9.94±14.55 (0.00-86.36) N=67	9.81±14.94 (0.00-86.36) N=73	NS
% uncited documents	Tenured Scientist	6.75±15.92 (0.00-100.00) N=67	6.06±10.75 (0.00-50.00) N=105	6.33±12.97 (0.00-100.00) N=172	NS
	Research Scientist	7.71±12.68 (0.00-42.11) N=30	7.30±14.54 (0.00-100.00) N=61	7.43±13.89 (0.00-100.00) N=91	NS
	Research Professor	8.33±20.41 (0.00-50.00) N=6	7.74±11.32 (0.00-75.00) N=67	7.79±12.10 (0.00-75.00) N=73	NS

Figure 10. Percentage of uncited documents of BIOL scientists by scientific category and gender



Scientific productivity and impact by number of years at CSIC and gender

After controlling for the number of years at the institution, an unequal distribution of men and women by scientific categories was observed (Figure 3). Especially for those who had an intermediate length of working life at CSIC (11 - 30 years), men were more likely to have been promoted than women. Could this fact be justified by higher productivity and/or impact of men as compared to women for these specific lengths of stay at the institution?

Table 3 shows inter-gender differences in productivity among scientists with less than 30 years at CSIC (p<0.05). Among the youngest scientists, with less than 10 years at CSIC, women are more productive than men (10 vs. 8 documents), while the opposite is found in the next two classes, in which men are more productive than women. Although men show slightly higher values of observed impact in all seniority classes, the differences between the sexes were not significant. However, it should be noted that the maximum values of productivity and citations belonged mostly to men.

Figure 11.
Productivity of BIOL scientists by years at CSIC

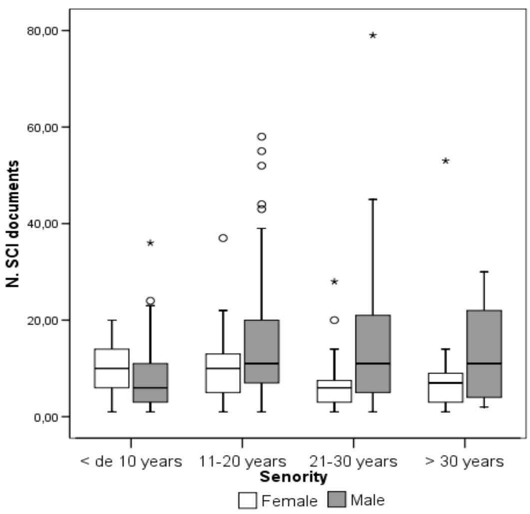


Figure 12.
international visibility of BIOL scientists by years at CSIC and gender

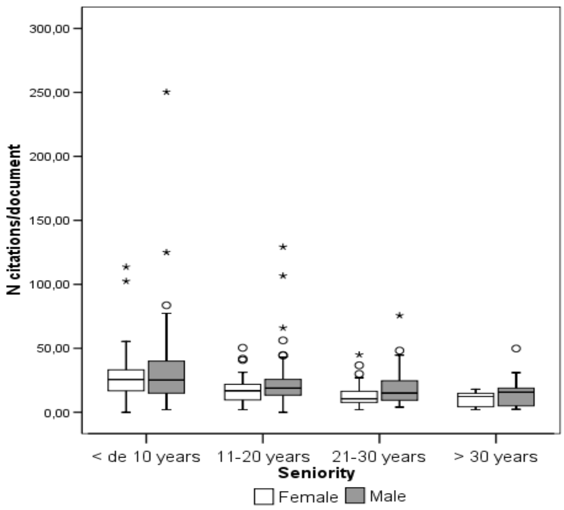


Figure 13. Percentage of highly cited documents of BIOL scientists by years at CSIC and gender

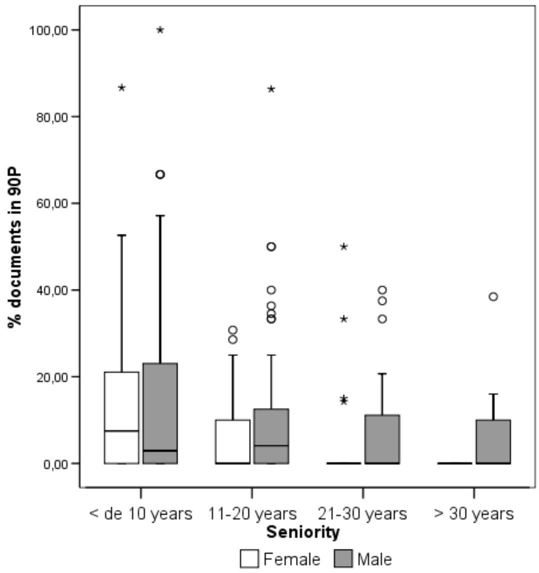


Figure 14. Percentage of uncited documents of BIOL scientists by years at CSIC and gender.

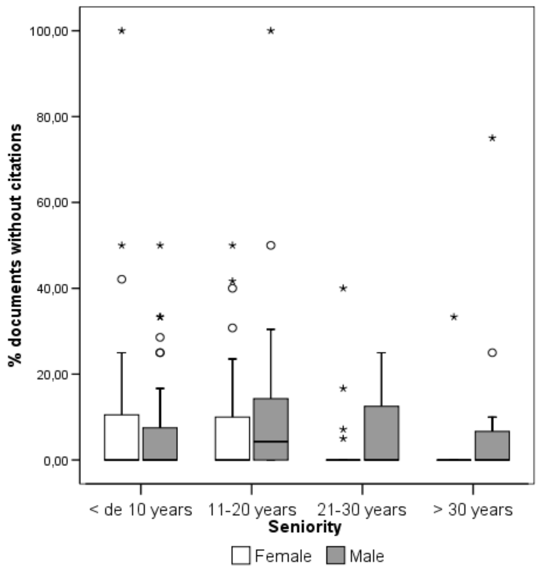


Table 3. Productivity and international visibility of BIOL scientists by years at CSIC and gender

		Female	Male	Total	SIG
No. Documents SCI	< 10 years	10.21±5.72 (1.00-20.00) N=34	7.99±6.81 (1.00-36.00) N=72	8.70±6.54 (1.00-36.00) N=106	0.05
	11-20 years	10.02±7.10 (1.00-37.00) N=41	15.10±11.99 (1.00-58.00) N=94	13.56±10.97 (1.00-58.00) N=135	0.05
	21-30 years	7.47±6.73 (1.00-28.00) N=19	14.54±13.74 (1.00-79.00) N=50	12.59±12.57 (1.00-79.00) N=69	0.05
	> 30 years	11.56±16.06 (1.00-53.00) N=9	13.12±9.64 (2.00-30.00) N=17	12.58±11.94 (1.00-53.00) N=26	NS
Av. IF	< 10 years	6.02±3.14 (2.72-16.44) N=34	6.51±3.44 (1.44-19.68) N=72	6.35±3.34 (1.44-19.68) N=106	NS
	11-20 years	4.74±2.04 (1.18-11.34) N=41	4.98±2.40 (0.69-14.00) N=94	4.91±2.29 (0.69-14.00) N=135	NS
	21-30 years	4.37±2.93 (0.92-14.00) N=19	4.19±2.60 (1.29-16.85) N=50	4.24±2.67 (0.92-16.85) N=69	NS
	> 30 years	3.36±2.25 (0.92-7.95) N=9	4.77±2.81 (1.31-13.63) N=17	4.28±2.67 (0.92-13.63) N=26	NS
No. citations/document	< 10 years	29.61±22.96 (0.00-113.63) N=34	33.00±33.84 (2.00-250.33) N=72	31.91±30.70 (0.00-250.33) N=106	NS
	11-20 years	17.26±10.97 (2.00-50.38) N=41	22.51±18.04 (0.00-129.23) N=94	20.91±16.36 (0.00-129.23) N=135	NS
	21-30 years	14.87±11.61 (2.00-45.00) N=19	18.44±13.30 (4.00-75.67) N=50	17.45±12.87 (2.00-75.67) N=69	NS
	> 30 years	9.93±6.11 (2.00-18.00) N=9	15.78±12.79 (2.33-49.85) N=17	13.75±11.16 (2.00-49.85) N=26	NS

		Female	Male	Total	SIG
% highly cited documents	< 10 years	12.77±18.52 (0.00-86.67) N=34	15.46±21.99 (0.00-100.00) N=72	14.60±20.89 (0.00-100.00) N=106	NS
	11-20 years	5.05±9.03 (0.00-30.77) N=41	8.90±13.87 (0.00-86.36) N=94	7.73±12.69 (0.00-86.36) N=135	0.05
	21-30 years	5.93±13.69 (0.00-50.00) N=19	6.86±10.02 (0.00-40.00) N=50	6.60±11.05 (0.00-50.00) N=69	NS
	> 30 years	0.00±0.00 (0.00-0.00) N=9	5.85±10.15 (0.00-38.46) N=17	3.82±8.60 (0.00-38.46) N=26	NS
% uncited documents	< 10 years	9.58±19.89 (0.00-100.00) N=34	5.30±10.37 (0.00-50.00) N=72	6.67±14.18 (0.00-100.00) N=106	NS
	11-20 years	7.46±13.41 (0.00-50.00) N=41	8.40±13.45 (0.00-100.00) N=94	8.12±13.40 (0.00-100.00) N=135	NS
	21-30 years	3.62±9.73 (0.00-40.00) N=19	5.86±7.77 (0.00-25.00) N=50	5.25±8.34 (0.00-40.00) N=69	NS
	> 30 years	3.70±11.11 (0.00-33.33) N=9	8.01±18.38 (0.00-75.00) N=17	6.52±16.13 (0.00-75.00) N=26	NS

Discussion

The purpose of this study is to analyse inter-gender differences in the scientific activity of CSIC scientists in the area of Biology and Biomedicine by means of bibliometric indicators. Our results show that the incorporation of the gender dimension into bibliometric studies is possible, although laborious due to the lack of standardisation of author names, and that it provides useful data for the comparison of the scientific activity of men and women.

<u>Methodological problems</u>

Obtaining bibliometric indicators by gender is not an easy task. Several problems arise:

1. Micro-level related problems. Studies at the micro level (individual level), such the one here developed, present special problems when compared with those of larger units of aggregation such as institutions, regions or countries. For example, statistical significance is harder to demonstrate when applied to small units and greater precision is required in terms of data collection, since small losses of information might have important influence on the final results.

2. Identification of the sex of the authors from publications. Very often the sex of an author can not be deduced from publications since only the initials of the first names are recorded. In this paper, the fact that we knew the subjects' sex avoided this problem. Obtaining bibliometric indicators by gender on specific institutions is thus an interesting alternative since the institutions can supply full information on the research personnel (complete name, sex, working centre, etc).

3. Lack of normalisation of author names. Our main challenge in this study was to identify correctly the scientific publications of each scientist. Due to the lack of standardisation of author names in publications, an author's production was potentially dispersed under a variety of signatures, and we had to normalise authors' names before obtaining bibliometric indicators at the micro level.

The indexing policies for author names vary between databases and over time within a specific database. At the time this research was carried out, the Thomson-ISI database only included the initial of the first name of the authors and not the full name. This database considered that the last part of the author's name was the surname, and included after a blank space the initial(s) of the first part that were considered as the first name. Our methodology took into account this indexing policy for the adequate identification of scientists in the by-line of publications.

Improving normalisation of authors names

With the aim of make future bibliometric studies by gender easier, authors, journals and databases should be aware of the importance of the standardisation of authors' names in publications. This will be positive for authors, since the use of standardised names will facilitate the correct identification of their production. For journals and databases, standardisation is also positive since it will contribute to increase the quality of these vehicles of information. In particular, we consider that

- The inclusion of the full name of the authors –and not only the initial of the first name- in the publications is essential. As a result, the problems of homonyms will be reduced and the correct identification of authors will be easier.

- Authors should select how they are going to sign in publications and maintain the signature all along their professional life. In this line, Ruiz-Pérez (2002) recommends establishing a "pen name", a signature name that it is best not to change over time and under which all works should be signed.

The problems derived from the lack of normalisation of authors names in publications have been analysed by different authors. The study conducted by Ruiz-Pérez et al. (2002) on the variability of Spanish authors' signatures in bibliographic databases clearly shows the magnitude of the problem and includes a series of recommendations to authors, journals and databases, which are the different actors who interplay in the process of publication and dissemination of knowledge. Journals and publishers can play a pivotal role in standardising author's signatures by asking authors to always sign their works using the same name, in the most complete possible form and including the workplace along with their signatures.

In our study, we have dealt with Spanish names whose structure is well-known by the authors of this paper. However, the structure of names is frequently linked to the cultural traditions of each country as well as the idiomatic variations (Borgman and Siegfried, 1992). As an example, the Anglo-Saxon world usually uses only one surname, whereas in the Hispanic world it is much more common to use two surnames. The inclusion of hyphens to join different particles of the first name (i.e. *Ana-Maria Martínez*) or different particles of the surnames (i.e. *Ana De-la-torre*) has been suggested as a way to avoid incorrect citations of authors by other scientists or by database indexers (Ruiz-Pérez et al, 2002). Standardisation of authors' names is essential to help us to cope with this inter-country variability.

One idea that is worth pursuing to address these problems is that of a Digital Author Identifier, analogous to a Digital Object Identifier, which would assign a unique number to a scholarly author the first time he or she published anything and would then be used consistently thereafter by that author.

Inter-gender comparison of scientific activity

Our results suggest a slower promotion for female scientists compared to their male counterparts. Scientific publications play an important role in the research process since they make the dissemination of the new knowledge possible. Scientific publications and citations are taken into account in the promotion of scientists, so inter-gender differences in productivity and impact could influence and contribute to explain differences in the promotion of female and male scientists. Are women less productive than men in the BIOL area at CSIC? Do women publish in less prestigious journals? Do they obtain fewer citations?

Productivity

In our study, there were no statistically significant differences in productivity of men and women. Since productivity increases with scientific category, the lower presence of women in the upper and most productive category might have some negative influence on the average productivity of women. To avoid this problem, the analysis of productivity by scientific categories is particularly appropriate for the comparison between genders (Bordons et al, 2003). After including scientific category in the analysis, again no statistically significant differences in productivity between men and women within each scientific category were found. But looking into productivity by years at the institution, we observed that women with less than 10 years at CSIC are more productive than men, whilst men are more productive than women in two seniority classes (11-20 years, 21-30 years at the institution). The last observation might contribute to explain the greater male promotion.

Lower productivity for female scientists has been observed in different studies (Cole and Zuckerman, 1984). This issue has been called the “productivity puzzle” (Cole and Zuckerman, 1984), still unresolved, and different underlying factors have been proposed. Among the factors suggested we can mention the lower trend of women to collaborate (Kyvik and Teigen, 1996), their lower social prestige, which limit access to economic and material resources (Xie and Shauman, 1998), their higher child care responsibilities (Kyvik and Teigen, 1996) and high teaching loads. The possible influence of personal factors has also frequently been pointed out and several authors have suggested that men are more ambitious and more secure as a result of educational and cultural factors; they are more aggressive and self-promoting in the pursuit of career success and these features can benefit them in the highly competitive science environment. On the other hand, women are less concerned with the political aspects of science, such as influence and power (Holton, 1999). Moreover, some authors consider that men and women differ in their “working styles” (Fox, 1999).

Impact

The impact factor of journals is very often used as an indicator of the expected impact of documents and as a proxy for quality, since within each discipline; journals with the highest impact factor are typically the most prestigious ones. On the other hand, citations are used as indicators of the observed impact of the research upon the scientific community. However, both indicators have limitations -widely described in the literature- and should be used with caution (Norris and Oppenheim, 2003, Moed, 2005).

In our study, no significant differences in the impact factor of journals or in the number of citations per document of male and female scientists were found. However, men performed better than women in two aspects: those with 11-20 years at CSIC showed a higher percentage of highly cited documents, and those in the lowest category show a significantly higher number of citations per document than women.

Although previous studies have suggested that women publish less documents than men but of higher quality (higher number of citations per document) (Feller, 2004; Long, 1992), our results do not support this hypothesis.

However, we should keep in mind that the number of citations received can be influenced by different factors such as: type of research (typically, higher number of citations for basic research than for applied research), collaboration (national but particularly international collaboration has been associated with higher impact) or research topic (hot topics attain higher levels of citations). Previous studies have shown less collaboration for women (Holton, 1999; Kyvik and Teigen, 1996), lower integration in social networks within the scientific community (Fox, 1999), and lower mobility and lower tendency to make overseas contacts through travel (Lewison 2001), factors which could influence negatively on the probability of women to be cited, whatever the quality of their research. Unfortunately, we do not have data to analyse the influence of these aspects over the performance of women in BIOL at the Spanish CSIC.

Slower promotion of women?

Although scientific publications are an important output of research activity, promotion does not rely only on publications. Other dimensions of research performance such as research projects, contracts, supervising PhD students or prizes and awards received are also taken into account.

But productivity increases with scientific category for male and female BIOL scientists, data that suggest that scientific publications are indeed a major factor in promotion decisions. Interestingly, impact as measured through the number of citations received does not tend to increase with category either for men or women. The high productivity of scientists in the top ranks has been frequently linked to their higher collaboration –frequently explained by their role as team leaders- but it does not imply a higher impact of research. Our results suggest that productivity is more important than impact for promotion.

Slower promotion for women can be deduced from our data, since for the same length of working life at the institution, women and men are not equally distributed by professional ranks. A relatively high percentage of women with 11-20 and 21-30 years at the institution remain in the lowest category, and the lower productivity of women as compared to men at these levels of seniority could contribute to the lower female promotion.

Nonetheless, women with less than 10 years at the institution are more productive than men in our study. Our data suggest that more competitive women have entered CSIC in the last 10 years. In that case, a more balanced distribution of women over scientific categories would be expected in a few years. However, maybe ten years ago, the youngest women were as productive –or more productive- than men and they changed their publication activity over the years as a response to social, economic or personal reasons. Unfortunately, we only have a static snapshot of the activity of men and women in a five-year period. Longitudinal studies would be necessary, to follow scientists during a large number of years and detect possible changes in their behaviour over time.

Conclusions

Our results show that bibliometric indicators by gender provide interesting data to analyse the scientific activity of women as compared with men. Obtaining these indicators is at present very laborious and time-consuming, but the development of a Digital Author Identifier would greatly help such studies in the future. The inclusion of different personal and social variables in these types of studies would be very desirable to increase our knowledge about the situation of women in research and detect possible barriers for their progression in their scientific careers.

References

- BIOSOFT SAS (2001), Development of Bibliometric Indicators by Gender. Final report under European Commission's contract ERBHPV2-CT-1999-14N.
- BORDONS, M., MORILLO, F., FERNANDEZ, M. T., GOMEZ, I. (2003), One step further in the production of bibliometric indicators at the micro level: Differences by gender and professional category of scientists, Scientometrics, vol. 57, nº 2, p. 159-173.
- BORGMAN, C. L.; SIEGFRIED, S.L. (1992). Getty´s Synoname and its cousins: a survey of applications on Personal Name-Matching Algorithms. Journal of the American Society for Information Science, vol. 43, nº 7, p. 459-476
- COLE, J. R., ZUCKERMAN, H. (1984), The productivity puzzle: persistence and change in patterns of publication of men and women scientists, Advances in Motivation and Achievement, 2 : 217-258.
- COSTAS, R.; BORDONS, M. (2007). Algoritmos para solventar la falta de normalización de nombres de autor en los estudios bibliométricos. Investigación bibliotecológica: archivonomía, bibliotecología e información vol. 21, nº 42, p. 13-32.
- COSTAS, R. IRIBARREN-MAESTRO, I. (2007), Variations in content and format of ISI databases in their different versions: The case of the Science Citation Index in CD-ROM and the Web of Science, Scientometrics, vol. 72, nº 2, p. 167-183
- CRANE, M. (2002). La place des femmes au CNRS: ce qui a changé depuis quinze ans. Centre National de la Recherche Scientifique. Direction des Etudes et des Programmes. Unité d´Indicateurs de Politique Scientifique. http://www.sg.cnrs.fr/ipam/publications/divers/place_des_femmes_au_CNRS.pdf (last accessed 25 July 2007)
- ETAN expert working group on Women and Science. Science Policy in the European Union. Promoting excellence through mainstreaming gender equality. European Commission. Research Directorate-General. Brussels, 2000. http://www.cordis.lu/improving/women/documents.htm (last accessed 25 July 2007)
- FELLER, I. Measurement of scientific performance and gender bias. In: Gender and Excellence in the making. Luxembourg: Office for Official Publications of the European Communities, 2004
- FOX, M.F. Gender, knowledge and scientific styles. Ann NY Acad Sci, Apr 1999, 869, 89-93.
- HEMLIN, S. and GUSTAFSSON, M. (1996), Research production in the arts and humanities: a questionnaire study of factors influencing research performance. Scientometrics, vol. 37, nº 3, p. 417-432
- GONZÁLEZ DUARTE, R. (Coord.) (2004). Documento sobre mujeres y ciencia. Observatori de Bioética.
- HOLTON, G. Different perceptions of "good science" and their effects on careers. Ann NY Acad Sci, Apr 1999, 78-86.
- Lawler, A. Universities urged to improve hiring and advancement of women. Science 313, 22 Sept 2006, 1712.
- KYVIK, S. and TEIGEN, M. (1996), Child care, research collaboration and gender differences in scientific productivity, Science Technology and Human Values, 21: 54-71.
- LEWISON, G. (2001), The quantity and quality of female researchers: a bibliometric study of Iceland. Scientometrics, 52: 29-43.

- LONG, J.S. (1992), Measures of sex differences in scientific productivity, Social Forces, vol. 71, nº 1, p. 159-178.
- MAULEÓN, E. AND BORDONS, M. Productivity, impact and publication habits by gender in the area of Materials Science. (2006). Scientometrics, 66: 199-218
- NORRIS, M.; OPPENHEIM, C. (2003). Citation counts and the Research Assessment Exercise V: Archaeology and 2001 RAE. Journal of Documentation, vol. 59, nº 6, p. 709-730
- MOED, H. F. (2005), Citation analysis in research evaluation, Springer, Berlin.
- PALOMBA, R. and MENNITI, A. (Eds) Minerva´s daughters. Institute for Research on Population and Social Polices. Rome, 2002
- PRAVDIC, N.; OLUIC-VUKOVI, V. (1991). Distribution of scientific productivity: ambiguities in the assignment of author rank. Scientometrics, vol. 20, nº 3, p. 131-144.
- PRPIC, K. (2002), Gender and productivity differentials in science, Scientometrics, 55: 27-58.
- RUIZ-PÉREZ, R.; DELGADO LÓPEZ-CÓZAR, D.; JIMÉNEZ-CONTRERAS, E. (2002). Spanish personal name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. Journal of Medical Library Association, vol. 90, nº 4, p. 411-430.
- SÁNCHEZ PEÑAS, C. and Willet, P. (2006). Brief Communication. Gender differences in publication and citation counts in librarianship and information science research. Journal of Information Science, 32: 480-485
- UIS Bulletin (2006). Women in science: under-represented and under-measured. UIS Bulletin on Science and Technology Statistics, 3, November 2006. (www.uis.unesco.org).
- ENWISE Report. Waste of Talents: turning private Struggles into a public Issue: Women and Science in the Enwise countries. European Commission. Directorate General for Research, Directorate C-Science and Society, Unit C4-Women and Science. Brussels, 2003
  http://ec.europa.eu/research/science-society/pdf/01_enwise_report_cover-140704.pdf
- WEBSTER, B. M. (2001), Polish women in science: a bibliometric analysis of Polish science and its publications. 1980–1999, Research Evaluation, 10 : 185-194.
- WIR Report. Women in Industrial Research: A wake up call for European Industry. STRATA-ETAN Report. European Commission. Directorate-General for Research. Brussels, 2003
  http://ec.europa.eu/research/science-society/women/wir/pdf/wir_final.pdf
  (last accessed 25 July 2007)
- She figures, 2006. Women in Science: statistics and indicators. European Commission. Directorate-General for Research, Directorate C-Science and Society, Unit C.4-Women and Science. Brussels, 2006 http://kif.nbi.dk/She_Figures_2006.pdf
  (last accessed 25 July 2007)
- XIE, Y., SHAUMAN, K. A. (1998), Sex differences in research productivity: new evidence about an old puzzle, American Sociological Review, 63, p. 847–870.

Acknowledgments

This study forms part of a research project supported by the Spanish *Instituto de la Mujer*. We are grateful to the Spanish *Ministerio de Educación y Ciencia* for the FPU fellowship conceded to Elba Mauleón and for her research stay in Loughborough.