Bibliometric Study of Scientific Research on Prion Diseases Encephalopathy and Creutzfeldt-Jakob Disease, 1973-2002

E. Sanz-Casado[*], C. Suárez-Balseiro[**], I. Iribarren-Maestro[*], M. Ramírez-de Santa Pau[***] and J. de Pedro-Cuesta[***]

[*]*elias@bib.uc3m.es, iiribarr@bib.uc3m.es*
Laboratorio de Estudios Métricos de Información (LEMI) - Departamento de Biblioteconomía y Documentación. Universidad Carlos III de Madrid. C/ Madrid 126, 28903, Getafe, Madrid

[**]*suarezbc03@yahoo.es*
Observatorio de Estudios Relacionados con la Informacion (OERI). Escuela Graduada de Ciencias y Tecnologias de la Información. Universidad de Puerto Rico, Recinto de Rio Piedras

[***]*mramirez@isciii.es, jpedro@isciii.es*
Instituto de Salud Carlos III. C/ Sinesio Delgado 6. 28029 Madrid.

Abstract
The purpose of the present study was to analyse the trends in research on prion diseases encephalopathy and Creutzfeldt-Jakob disease by applying bibliometric tools to the scientific literature published between 1973 and 2002. The data for the study were obtained from Medline database. The aim was to determine the volume of scientific output in the above period, the countries involved and the trends in the subject matters addressed. Significant growth was observed in scientific production since 1991 and particularly in the 1996-2001 period. The countries found to have the highest output were the United States, the United Kingdom, Japan, France and Germany. The collaboration networks established by scientists were also analysed in this study, as well as the evolution in the subject matters where they were being published, that were almost constant in the three sub-periods in which the study was divided.

Introduction
Bovine spongiform encephalopathy (BSE), a new epidemic essentially affecting bovine cattle, was first detected and diagnosed in the United Kingdom in 1986. It has since been described in many countries, primarily in Europe. The first report of animal spongiform encephalopathy dates from the eighteenth century; known as scrapie, an endemic disease in sheep with a high rate of occurrence in Iceland, its transmission was proven experimentally in 1936 (Brown, Bradley & Cathala, 1999).

In the nineteen twenties, Creutzfeldt and Jakob described a sub-acute or chronic encephalopathic process in patients (Escudero-Torrela, 2000) occurring in the population at large at a rate of approximately one per million. The syndrome began to be studied in greater depth essentially after Gajdusek's findings on kuru in New Guinea. The discovery in the sixties that the lesions caused by scrapie were similar to those found in kuru and the Creutzfeldt-Jakob syndrome (CJS) led to rapid progress in the molecular characterisation of the agents involved and their relationship with the existence of prions (Brown, Bradley & Cathala, 1999). Epidemiological data were obtained in 1996, followed by experimental data that established the connection between BSE and the so-called variant Creutzfeldt-Jakob disease (vCJS), transmissible to human beings via infectivity in food (Gargani, 2002).

The purpose of this paper is to analyse BSE and CJS research trends by applying bibliometric tools to the scientific literature published between 1973 and 2002 to determine the volume of scientific production, the countries involved and the trends in the research subjects addressed.

Methodology
The period analysed, running from 1973 to 2002, was divided into three sub-periods (1973-1982, 1983-1992, 1993-2002).

Proceedings of ISSI 2005, pages 595–603.
Edited by P. Ingwersen and B. Larsen.

The data for the study were obtained from Medline, the U.S. National Library of Medicine's (NLM) on-line PubMed service (http://www.ncbi.nlm.nih.gov). The information was retrieved from the NLM controlled vocabulary database (MeSH) of PubMed contents; the query terms used were: encephalopathy, bovine spongiform [MESH] OR creutzfeldt-jakob syndrome [MESH] OR gerstmann-straussler-scheinker disease[MESH] OR insomnia, fatal familial[MESH] OR kuru[MESH] OR scrapie[MESH] OR prion diseases[MESH] OR prion[MESH] OR prion protein[MESH]

This strategy yielded 7,808 records. Some of these records proved to be duplicated or incomplete and were deleted. The final database consisted of 7,800 records, on which an analysis of all the information of bibliometric interest was conducted. Where no author affiliation information was available, the paper was looked up in the original source.

Both uni-dimensional (scientific production, co-authorship index, publication subject matter) and multidimensional (analysis of co-occurrence of terms) indicators were used; the latter as a technique to explore and study the inter-relationships among authors found in the papers reviewed (Sanz-Casado & Martín-Moreno, 1997). Multivariate analysis - multidimensional scaling and correspondence factorial analysis – was used to obtain these indicators and explore both author's collaboration networks and trends in the subjects researched during the sub-periods under consideration.

In the study on the relationships between subject categories, only the terms specified in the MeSH database as indicative of the main content of the paper were adopted. In order to make the study more comprehensible, these terms were standardised and in some cases aggregated along the lines suggested by the consulted BSE experts.

All statistical analyses were run with Microsoft Excel® and version 10.1 of SPSS® for Windows. The working matrices were built with Bibexcel® for Windows, an in-house software for bibliometrics analysis developed by professor Olle Persson and the Inforsk Research Group at Umea University, Sweden. Data were processed and indicators obtained in accordance with the methodology proposed for information metrics studies in a paper by Sanz et al (2002).

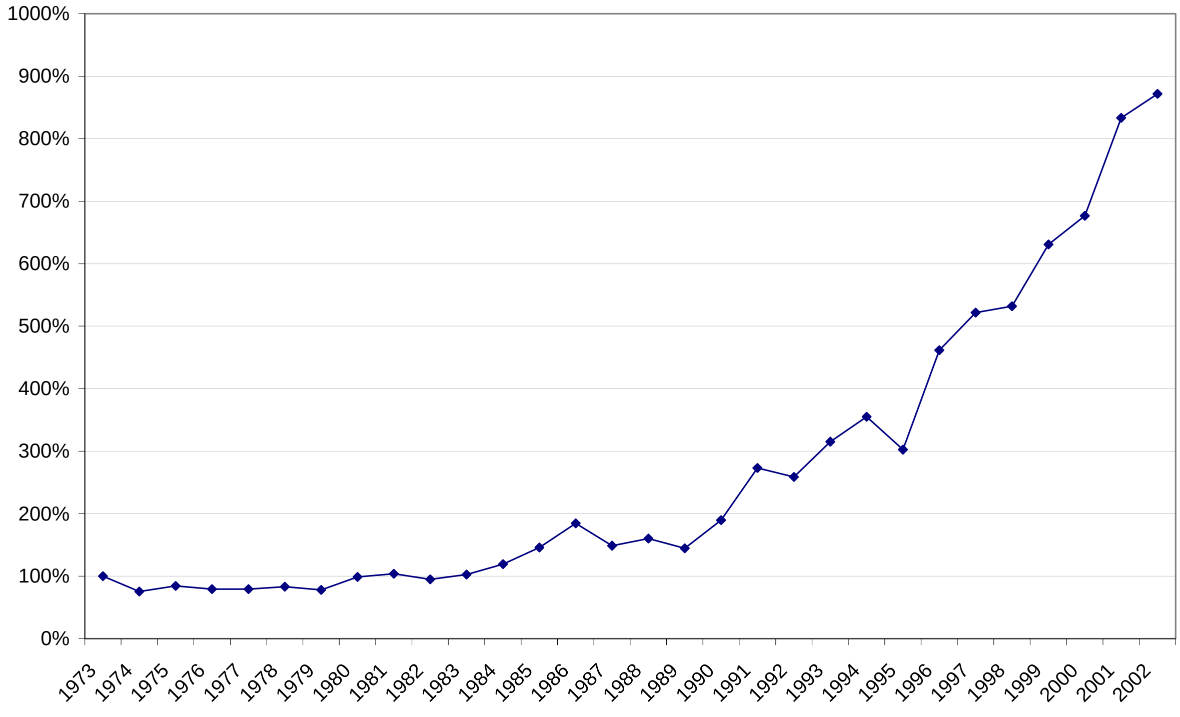

Figure 1. Trends in scientific production, 1973-2002

Results

Scientific production 1973-2002

Figure 1 shows the variations in scientific production between 1973 and 2002, in comparison with the 7,800 papers published and the annual increase with regard to 1973. A gradual increase in scientific production was observed from the seventies to the mid-eighties, with a steep rise at the end of that decade. Production increased very fast from 1990, with a peak in 1991, until 1996 where another sharp rise was recorded, with peaks in 1996. The number of publications grew by 94.87% between 1973 and 1982, and by 258.97% and 871.79% from 1973 to 1992 and 1973 to 2002, respectively.

Geographic distribution of scientific production

Table 1 shows the geographic distribution of scientific production in the subject areas in question, along with growth in absolute values and percentages in the sub-periods established. Only the percentages corresponding to countries accounting for over 1% of the papers published in each sub-period were considered. The country with highest scientific output throughout the period studied was the United States, with 40.19% and 41.08% of the papers published in the earlier decades, and 26.73% in the third sub-period. The second ranking country by number of papers published in the three periods was the United Kingdom (UK), with a ratio average higher than 20%. France and Japan presented also an important research activity. Whilst Japan's relative weight declined in the third sub-period, its publishing activity did not, as it was 2.4-fold higher in absolute terms than in the second decade.

Table 1. Breakdown of scientific production by country

1973-1982			**1983-1992**			**1993-2002**		
Country Name	No. pap. publ.	%	Country Name	No. pap. publ.	%	Country Name	No. pap. publ.	%
USA	211	40,19%	USA	463	41,08%	USA	1087	26,73%
UK	115	21,90%	UK	209	18,54%	UK	827	20,34%
France	38	7,24%	Japan	115	10,20%	France	394	9,69%
Japan	38	7,24%	France	57	5,06%	Germany	325	7,99%
Italy	14	2,67%	Germany	55	4,88%	Japan	286	7,03%
Australia	12	2,29%	Italy	43	3,82%	Switzerland	187	4,60%
Czechoslovakia	12	2,29%	Poland	31	2,75%	Italy	155	3,81%
Germany	9	1,71%	Russia	19	1,69%	Spain	74	1,82%
Switzerland	9	1,71%	Canada	13	1,15%	Holland	72	1,77%
Canada	8	1,52%	Spain	13	1,15%	Poland	72	1,77%
Poland	7	1,33%	Israel	13	1,15%	Canada	59	1,45%
Chile	6	1,14%	Switzerland	12	1,06%	Austria	58	1,43%
Spain	6	1,14%	Other*	84	7,45%	Australia	55	1,35%
Other*	40	7,62%		1127	100,00%	Israel	54	1,33%
	525	100,00%				Belgium	42	1,03%
						Other*	319	7,85%
							4066	100,00%

* Countries whose contribution accounted for less than 1% of the total

Scientific collaboration

Scientific collaboration has been established by means of co-authorship index which shows the size of research groups and the strong of links between them. Figure 2 shows the evolution of co-authorship index, which began at a very low value (2.3 authors per document). However, the values show an important growth from 1981 to 2002 reaching 4.14 authors per paper. Although we cannot identify a regular pattern in the evolution of co-authorship indexes, their values are higher than the average from 1981 to 2002.

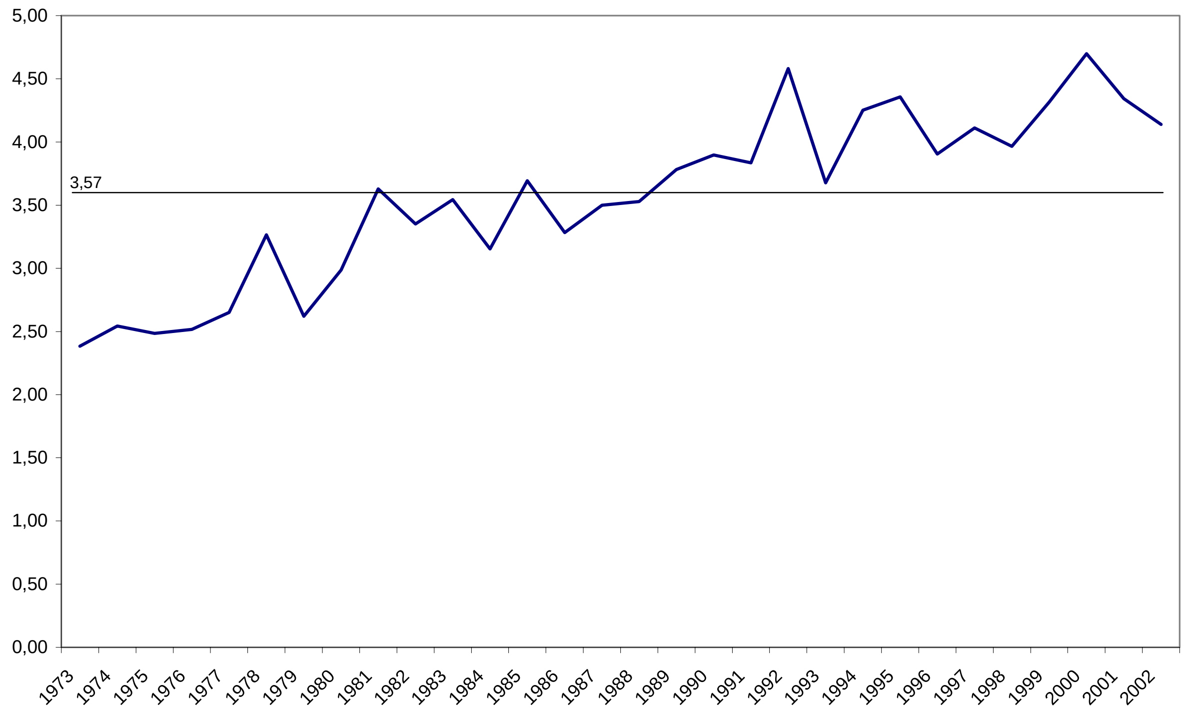

Figure 2. Evolution of the co-authorship index

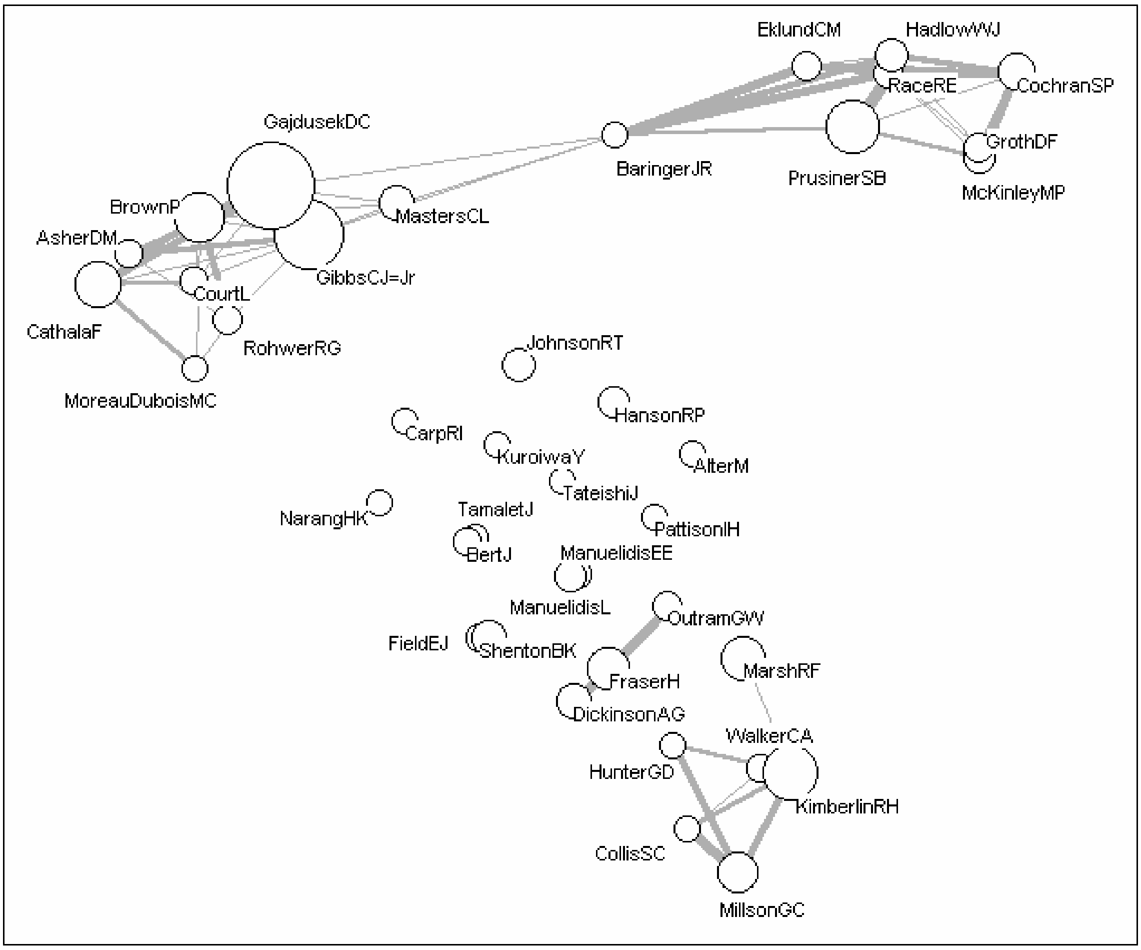

Figure 3. Collaboration networks during the 1973 – 1982 sub-period

Figures 3, 4 and 5 show the result of running multidimensional scaling technique to mapping collaboration networks in the three sub-periods established. We have considered only more productive authors (higher than 1% with regard to all papers in each period).

The map for the first sub-period (1973-1982) shows three well-defined clusters: the biggest, in the top left corner, is formed by: Gajdusek DC, Brown P, Cathala F, Gibbs CJJr, Asher DM, Court L, Masters CL, Moreau Dubois MC y Rohwer RG. The first four authors in this group, especially Gajdusek, are the most active one. A second cluster is in the top right corner and it is formed by: Prusiner SB, Cochran SP, Groth DF, McKinley MP and Baringer JR. In this case Prusiner is the most productive author. Note the position of Baringer which seems a connection between Gajdusek's cluster and Prusiner's cluster. The third cluster is situated in the bottom right zone of the map. It is formed by: Kimberlin RH, Millson GC, Marsh RF, Hunter GD, Collis SC, Hanson RP and Walker CA. Other clusters on this map are shorter than the three above-mentioned ones. Note that many of them have only two authors.

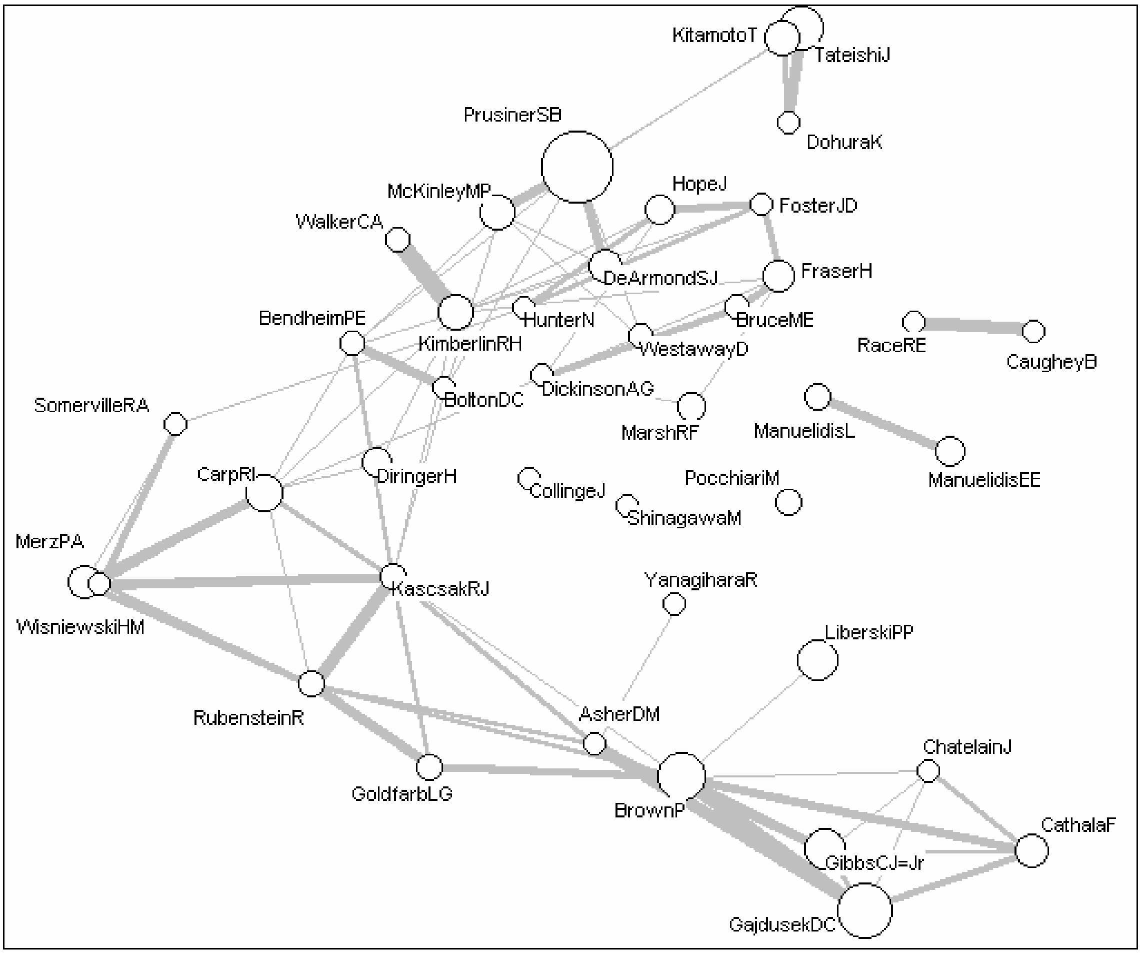

Figure 4. Collaboration networks during the 1983 – 1992 sub-period

The map for the second sub-period (1983-1992) shows a more complex network (Figure 4). There are more collaboration links for this decade and some changes can be observed with regard to the period between 1973 and 1982. The first cluster, in the bottom right of the map is the most productive. It is formed by: Gajdusek DC, Gibbs CJJr, Cathala F, Brown P, Chatelain J, Asher DM, Goldfarb LG, Liberski PP, Pocchiari M and Yanagihara R. The second cluster situated at the top of the map is formed by: Prusiner SB, Bendheim PE, DeArmond SJ, McKinley MP, Bolton DC, and Westaway D. Note that the group of Manuelidis EE and Manuelidis L, is isolated from the rest; in this sense these two authors show a similar behaviour with regard to collaboration links in the first sub-period analysed. In addition, there are two new clusters in this period, the first one formed at the top of the map by: CarpRI, Diringer H, Kascsak RJ, Kimberlin RH, Wisniewski HM, Rubenstein R, Walter CA, Merz PA, Kimberlin RH and Walter CA, and the second one formed by: Hope J, Foster JD, Dickinson AG, Hunter N, Marsh RF and Somerville RA.

The configuration of the last sub-period (1993-2002) is showed in figure 5. We can see important changes with regard to the size of groups and the links among them. A more concentrated network is the result of a more productive and collaborative environment. Note the growing number of clusters (12) and the increase in the "web" of links. Authors like Gajdusek, Brown and Prusiner have maintained high productive profiles and could be named "research leaders" for the three periods at all, although it is evident that they have diversified their contacts and partners for publications.

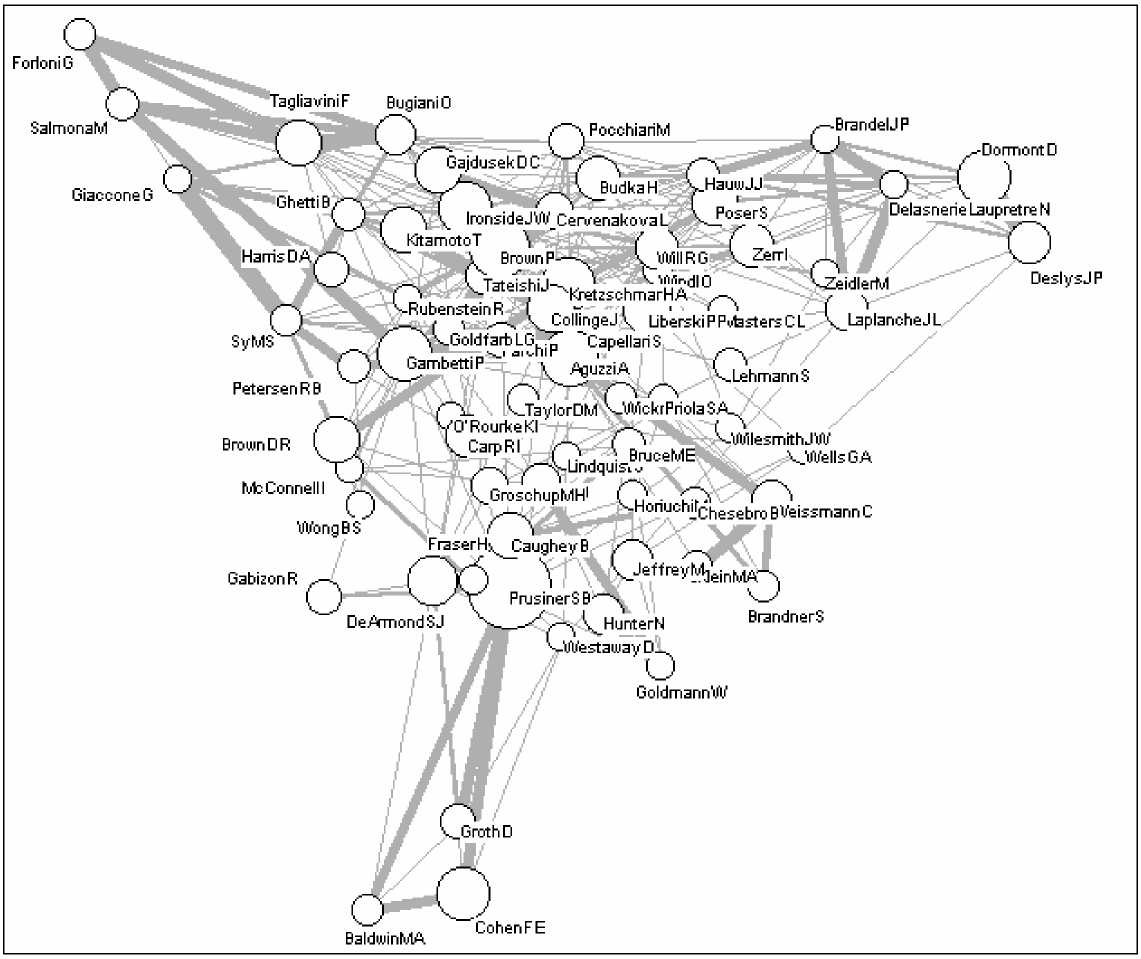

Figure 5. Collaboration networks during the 1993 – 2002 sub-period

Research trends: relation between clusters of authors and subject categories.
Figures 6, 7 and 8 show the distribution of author's clusters with regard to a group of research subjects-matters for each sub-period. Correspondence Factor Analysis has been used to map the relation between these two categories.

Figure 6 shows the picture for the 1973-1982 sub-period. We can appreciate some clusters close around the coordinate's origin among some high productive subject-matters (i.e. scrapie, nervous system, chemicals and drugs, diseases). Note the position of cluster 1 (C1) which seem the most productive in many subjects. The position of cluster 5 (C5) suggest it is focused on Anatomy, Kuru-immunology and CJS-immunology. At the centre left is the cluster 2 (C2) which seem more active in Scrapie, Chemicals and Drugs and Nervous-Systems-Diseases. Cluster 4 (C4) seems more focused on Scrapie. Note the position of clusters 7 (C7) formed by the two isolated above-mentioned authors (Manuelidis EE and Manuelidis, L.) which appear nearest to Psychiatry & Psychology and CJS-Pathology.

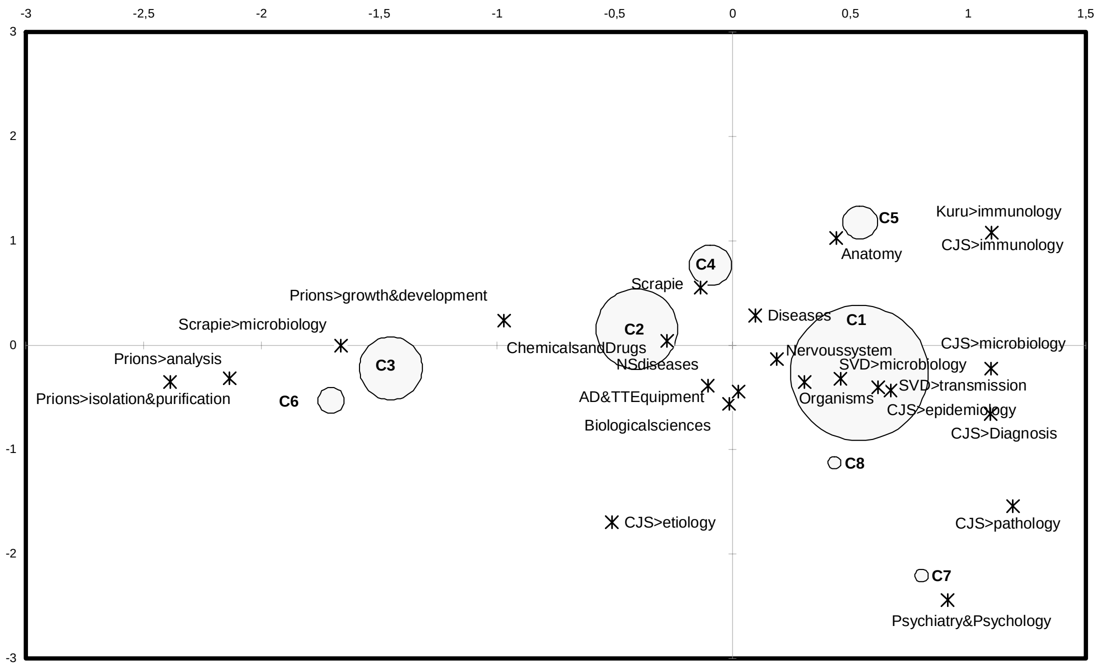

Figure 6. Author's clusters vs research subjects-matters (1973-1982 sub-period)

Figure 7 shows the picture for the second sub-period (1983-1992). An important fact is the change in position with regard to the first sub-period. Most of clusters are moving towards a central zone. It could be a clue that authors are being more active in a group of subject matters. Cluster 1 (C1) is the most productive. Nearest to C1 is C5 formed by Japanese researchers (Tateishi, Kitamoto, Dohura) which are related to CJS-Pathology. Cluster 2 (C2) situated at the top side seems to be related to Prions-analysis, Prions-genetics, Prions genetics and Purification. Note that cluster 4 (C4) has a peripheral position, and it seems more related to subjects like CJS-metabolism.

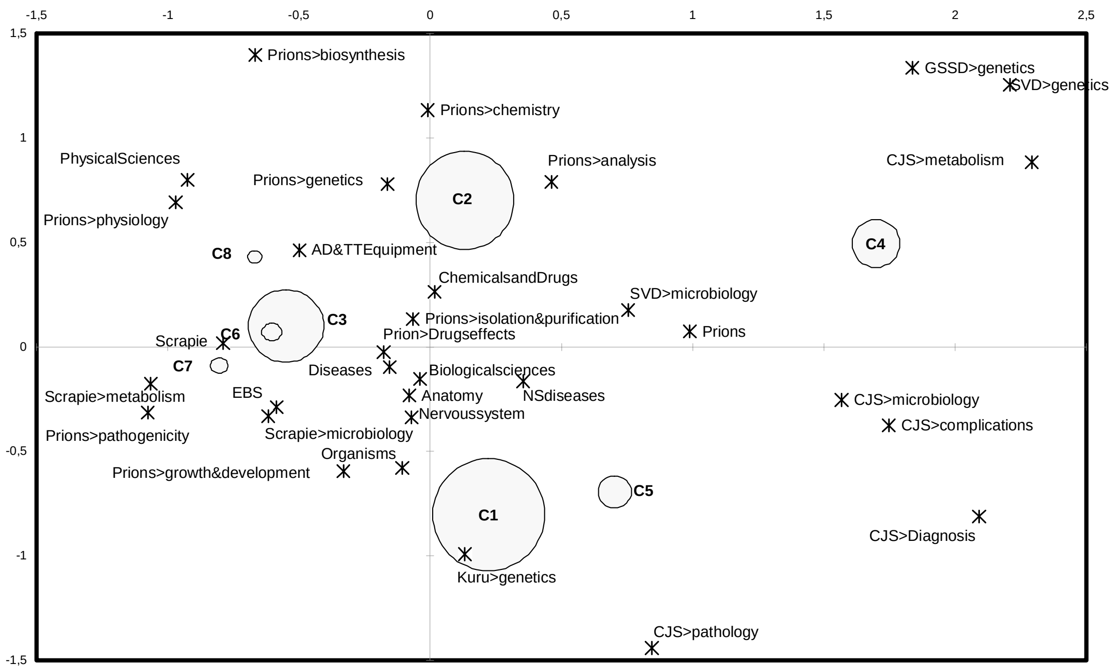

Figure 7. Author's clusters vs research subjects-matters (1983-1992 sub-period)

Figure 8 shows the situation for the last sub-period (1993-2002). We put an emphasis on this more complex picture, with a lot of clusters and subject matters very close to the barycentre. Subjects like: Chemical and drugs, Nervous-system, Biological-sciences, Diseases, Nervous-system-diseases and anatomy, seem to define the dominant trends on the map. Also clusters 1, 5 and 4 appear at the right side and very close to each other. It suggests similarities in subject research profiles. At the right side are two clusters (C6 and C11) which show a publication pattern related to some subjects in the central zone but a focus on CJS-diagnosis, CJS-epidemiology, CJS-pathology, and others. Close to the centre are three clusters (C8, C3, and C9). The profiles of these clusters are related to subjects placed in the centre but also to other topics like Prion-disease-genetics, Prions-pharmacology, and Scrapie. Finally, shortest clusters appear at the bottom left of the map (C10 and C12). These clusters seem to be related to Scrapie and EBS.

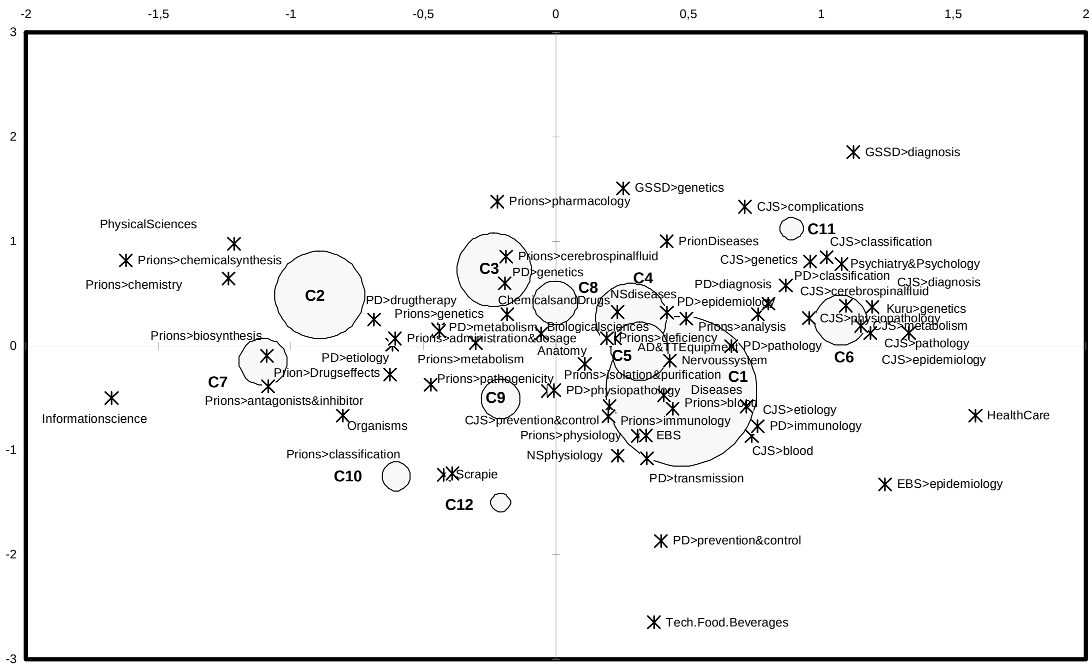

Figure 8. Author's clusters vs research subjects-matters (sub-period 1993-2002)

Discussion and conclusions

The most notable findings in the study include the fast increase of the number of publications on these subjects after 1990. With regard to scientific activity by geographic area, greater output in other countries was responsible for the relative decline in US production. Scientific output in the United States is characterised by intense research in nearly all areas of knowledge, possibly spurred in this case by the impact of Gajdusek's research on kuru and Prusiner's studies on prions. The United Kingdom maintains the second place by number of papers published in the three sub-periods because its output was substantially increased in the last sub-period as a result of the effect of BSE; and nearly all the victims of variant Creutzfeldt-Jakob syndrome were diagnosed in that country, with only approximately 5% confirmed outside the British Islands. A similar pattern was observed in Switzerland, where several cases of BSE also appeared earlier on. The rise in scientific production in countries such as Germany or Spain may have been due to the existence of CJS registries and research groups associated with diagnosis and epidemiological monitoring.

The values of co-authorship index show that scientific collaboration between the researchers has had grown significantly from 1981 to 2002 reaching 4.14 authors per paper. The collaboration networks between authors in the three established sub-periods have been very important. The map for the first sub-period (1973-1982) shows three well-defined clusters, while the map for the second sub-period (1983-1992) shows a more complex network (Figure 4). There are more collaboration links for this decade and some changes can be observed with regard to the first one. The configuration of the last sub-period (1993-2002) shows important changes with regard to the size of groups and the links among them. A more concentrated network is the result of a more productive and collaborative environment.

In relation with the subject matters that make up the clusters the authors are working on, in the first sub-period, it can be observed that there are various clusters in the central zone of the map, and that the subject matters where more publications have taken place are Scrapie, Nervous system, Chemical and drugs or Diseases, while in the second sub-period, there is a migration of the majority of the clusters towards the central zone of the map, which indicates that there is a major coincidence in the research subjects in which they are published. Some of those subjects are: Nervous-systems, Chemical and drugs, Diseases, Nervous-systems-diseases or Biological-sciences.

In the last sub-period, it can be observed that both the number of participating clusters and the subject matters where they are published are increasing. It is also worth mentioning that, as in the previous period, the majority of the clusters occupy a central position, because they share many of the research subjects, which are similar to those of the previous periods (Chemical and drugs, Nervous-systems, Biological-sciences, Diseases, Nervous-systems-diseases o Anatomy).

Finally, attention is drawn to the usefulness of bibliometric tools in the analysis of research activity, for the precise insight they provide into the characteristics and evaluation of scientific output. The information obtained from applying these techniques, in conjunction with expert opinion, can provide valuable support for enhanced scientific policy decision-making. In this study the results obtained were presented, discussed and validated with researchers in this field, they are working in the Instituto de Salud Carlos III, research public centre integrated in the National Health System.

Acknowledgments

The authors of this study would like to thank Professor Ingwersen and his group for their advice and orientations and for teaching us the use of the Bibexcel programme.

References

Brown, P., Bradley, R., & Cathala, F. (1999). Brief history of transmissible spongiform encephalopathies. *La Revue du Practicien*, 49 (9):928-33.

Escudero-Torrela, J. (2000). Cronología de la nueva variante de la Creutzfeldt-Jakob disease. *Revista de Neurología*; 31(2):141-47.

Gargani, G. (2002). Transmissible spongiform encephalopathies. History, epidemiology, etiological, hyphotheses. *Minerva Médica*, 93(1):59-73.

Sanz Casado, E. & Martín Moreno, C. (1997). Técnicas bibliométricas aplicadas a los estudios de usuarios. *Revista General de Información y Documentación*, 7(2):41-68.

Sanz-Casado, E., Suárez-Balseiro, C., García-Zorita, C., Martín-Moreno, C. & Lascurain-Sánchez, Mª L. (2002). Metric Studies of Information: Approach to a Practical Teaching Methodology for this Type of Studies. *Education for Information*, 20(2):133-144.