Stem cell research. Bibliometric analysis of main research areas through KeyWords Plus

Abstract

Purpose – Research with stem cells is a biomedical venture with great scientific impact, and whose development flows over into many other areas. This article presents a dual analysis of Spain´s scientific output in this field during the period 1997-2007.

Design/methodology/approach – We used, on the one hand, bibliometric indicators of a basic nature, and on the other hand, techniques for the visualization and analysis of networks of scientific information based on a study of KeyWords Plus.

Findings – The Output is mainly concentrated in Cataluña and Madrid, and hospitals are the most productive centres (followed by health institutes), where the main authors are affiliated. Main categories are Hematology, Oncology and Biophysics. The outstanding areas of study we identified revolve around the therapeutic use of transplant of hematopoietic progenitors, the processes of generation, proliferation and differentiation of lines of cells, and the study of neurosciences.

Originality/value – This study provides an overview of Spanish research involving stem cells, detecting and representing the main areas of research. We consider the potential of KeyWords Plus in combination with the proposed methodology as particularly useful for the analysis and delimitation of a scientific domain.

Keywords Stem-cells, Spain, Co-word Analysis, KeyWords Plus, Bibliometric Analysis, Information Visualization

Paper type Research paper

1. Introduction

Research into stem cells is a biomedical field of great expectations. It was in 1949 when scientist J. Hammond (1949) discovered the method to maintain mouse embryos in culture *in vitro*. Since, research with stem cells has developed to the point where it holds very hopeful perspectives for the treatment of thus far incurable diseases. At present, research is mainly oriented toward developing new therapies for hematological, cardiovascular, neurodegenerative and genetic diseases, as well as cancer and diabetes, among others (Martínez Serrano and Bjorklund, 1996; Bishop *et al.*, 2002; Cao *et al.*, 2002; Di Giorgio *et al.*, 2007). Its intrinsic nature makes stem cell research transcend to other fields as diverse as politics, ethics, culture, and law, placing it in the arena of social controversy.

The use of bibliometric studies to comprehend and analyse scientific domains (Hjørland and Albrechtsen, 1995), together with the development and fine-tuning of new techniques and tools, facilitates decision-making in areas of scientific policy and reflects the "state of the art" of research at a given time. These processes, necessary for the evaluation of science (Camí *et al.*, 1997; Bordons and Zulueta, 1999) are a responsibility that no country can elude (Krauskopf, 2000) given the evident connections between advancement through research activity, economic growth and progress, and the enhanced well-being of society (Chinchilla-Rodríguez and Moya Anegón, 2007).

The number of scientific disciplines interrelated by stem cell research lends it an interesting yet complicated character (Zhao and Strotmann, 2011). Its interdisciplinarity presents a great challenge when delimiting and analyzing its thematic composition, demanding a very precise analysis. Precisely to face this challenge, bibliometrics has complementary tools that more recently include social network analysis (Perianes-Rodríguez, *et al.*, 2011) and the visualization of scientific domains (Wasserman and Faust, 1998; Boyack *et al.*, 2009; Leydesdorff and Rafols, 2009; Rafols *et al.*, 2010; Aharony, 2010; Chinchilla-Rodríguez *et al.*, 2010; Vargas-Quesada *et al.*, 2010).

When searching for reliable information, studies of this nature have traditionally used bibliographic databases as the most usual source (Fernández *et al.,* 1993). Though not designed for this purpose, the information they offer has demonstrated its bibliometric potential. However, they present limitations when resolving specific matters, for example the precise definition and delimitation of a topic. It is therefore essential to reveal the semantic structure established among documents by means of the bibliographic information contained in a database.

Overcoming such limitations calls for analytical methods that will allow us to arrive at the ideas and concepts that ultimately sustain the scientific discourse expressed in the literature. To this end, units of analysis smaller than thematic categories or journals may be used: namely, descriptors. We opted to use KeyWords Plus (KW+) based on the hypothesis that they could better reflect the conceptual essence of each document. These descriptors are automatically generated by the SCI from an algorithm that extracts key words from all the titles referenced or cited in the documents (Garfield, 1990, 1993).

2. Antecedents

The application and development of this type of analysis is nothing new. In the 1980´s there were attempts to reveal the structure of science and its evolution through word co-occurrence. This marked a relationship between the concepts that documents represent (Cambrosio et al. 1993, Courtial et al. 1994) and made manifest the structure and trends of a scientific discipline in view of the strength of associations among the representative terms in the published literature (Ding et al., 2001).

Authors such as Van Raan and Tijssen (1993) applied it to research on neural networks, concluding that there is an epistemological value that allows us to discover unsuspected relations among the concepts of a discipline besides revealing problems that otherwise might go undetected. Ding, Chowdhurry and Foo (2001) used it to map the intellectual structure of the field of information retrieval, showing patterns and trends within. Onyancha and Ocholla, in 2005, used it to help researchers and project managers identify new research lines in HIV/AIDS, as well as study the links established, to better plan research and formulate adequate scientific policies. Bessalar and Heimeriks (2006) used it to study the publications in Information Science that came out between 1986 and 2002 by means of the most relevant words in titles and references. Chalík and Jirina (2006) related it with Price´s cumulative law of science as applied to Economics. In turn, Lee (2008) used this model to identify trends and underlying aspects in the area of Information Security, and Neff and Corley (2009) applied it to the study of publications in the area of Ecology from 1970 to 2005. More recently, Leydesdorff *(in press)* focuses on co-word analysis in relation with the semantic measures of similarity patterns (correlations) and latent variables (analysis factor) using computational and statistical techniques. And Zulueta et al. (*in press*) use it to study publications involving health and women as recorded in Medline from 1965 to 2005, combining network visualization techniques and factor analysis.

Despite such a proliferation of studies, and such a broad area of application, there is very little evidence of preceding studies that resort to KW+ as units of analysis. Qin (2000) uses them to compare differences among the descriptors of the SCI obtained by automated indexing and the MeSH terms obtained by manual indexing using a controlled language. In later works (Chiau and Ho, 2007; Ho, 2007), a brief recount of the Author KeyWords (AKW) is incorporated as a complement to bibliometric analysis, overlooking KW+. Authors Li et al.(2009) complement their bibliometric analysis with AKW as well as KW+, whereas Su and Lee (2010) analyse documents pertaining to the field of “Technological Prospective” registered in the Web of Science through a topographic representation of AKWs alone.

Aside from the fact that there is scant reference to bibliometric uses of KW+, there are likewise few publications that systematically undertake analysis of the topic at hand, stem cell research. From the bibliometric standpoint, Ho et al. (2003)**,** tried to gauge the scientific output of Asia, focusing on Hong Kong, Singapore, South Korea and Taiwan. Li et al. (2009) carried out a more recent and comprehensive analysis, appraising world output from 1991 to 2006. The methodological approach of Leydesdorf (2005; 2006) involved using the term “stem cell” in a co-word study as an element of analysis to show differences in its use and measure its significance in diverse contexts.

3. Research Questions

This paper aims to answer a number of questions about Spanish research surrounding Stem Cells:

RQ1. What is the overall structure of stem cell research field? Does this structure allow us to characterize the field or extract some sort of conclusion from a thematic standpoint?

RQ2. Do co-word studies applied to categories supply complementary information? Are they suitable for fine-grained studies such as this? Could they be considered adequate units of analysis for thematic delimitation at the document level?

RQ3 Are Kw+ better measures of analysis than subject categories? Do the two provide the same information? Can they be considered adequate units of analysis for the thematic delimitation at the document level?

4. Material and methods

4.1 Sources

The source of bibliographic information selected for this study was the Science Citation Index (SCI) database of the Thomson Reuters Institute (2010) for Scientific Information (ISI).

4.2 Data processing

The search strategy consisted of using the term *stem cell* in the field *topic*, delimited by "spain" in the field *address* and parameterized by the time period [1997-2007]. The search in the field *topic* allowed us to locate terms or phrases in specific parts of the document such as title, abstract, AKW and KW+. In April 2008 we finished downloading the records into an *ad hoc* relational database. The total number of documents retrieved was 2,467.

The ISI database presents certain disadvantages for bibliometric applications. For this reason it was necessary to standardize and purge the data. Ninety-eight documents were eliminated for one reason or another —the date of publication did not coincide with our period of study, we detected documents mistakenly attributed to the domain of Spain by homonymy with the name of the institution, there were some cases of record duplication, etc. Therefore, the final number of documents used for this study was 2,369. Standardization was carried out manually for the authors as well as the Spanish addresses.

	Not Standardized	Standardized
Authors	7,809	7,397
Spanish address	5,121	4,935

Table 1. Standardization of data

4.3 Analysis and representation of information

In order to analyse the studies about stem cells from different perspectives, we adopted two separate approaches. First we carried out a basic bibliometric study of a descriptive nature, through which we could generally appraise the state of research in the area and characterize the main actors at meso and micro levels. Secondly, we undertook a thematic delimitation by means of distributive and network analysis, using as units of analysis the Journal Citation Report categories and the KW+ assigned to each document. We adopted co-occurrence as the unit of measure, given that it is widely accepted for obtaining relational information from the documents within a particular domain. The final result was a square matrix of N x N elements, where N stands for the units of analysis to be represented. For visualization and analysis of the data, two programs were used: Pajek (Batagelj, 2010) and VOSviewer (Van Eck and Waltman, 2010).

5. Results and Discussion

Below we present our results according to the type of study applied: the general data obtained by bibliometric analysis, and the results related with the thematic delimitation obtained by analyzing the categories and KW+. In either case, the body of documents we worked with was the same, numbering 2,369.

5.1 Bibliometric analysis

5.1.1 Output trends

There is a clear rise in scientific output involving stem cell research over the period studied (Figure 1). The thresholds of production we encountered are 94 documents for the year 1997, and 395 for 2007, meaning a gross increase of over 300% in the decade. Noteworthy is the period from the year 2000 to 2003, during which output stabilized. The year 1999 shows the highest rate of production, and 2003 would be the lowest. Such fluctuations in national output could be affected by a number of factors such as the amount of congresses held, technological advances, levels of international collaboration, and political policy or specific measures applied to the area of stem cell research. While true that we do not have all the information needed to identify the causes for this output trend, some data and observations could shed light on the aforementioned factors and therefore on the evolution of scientific production related to stem cells.

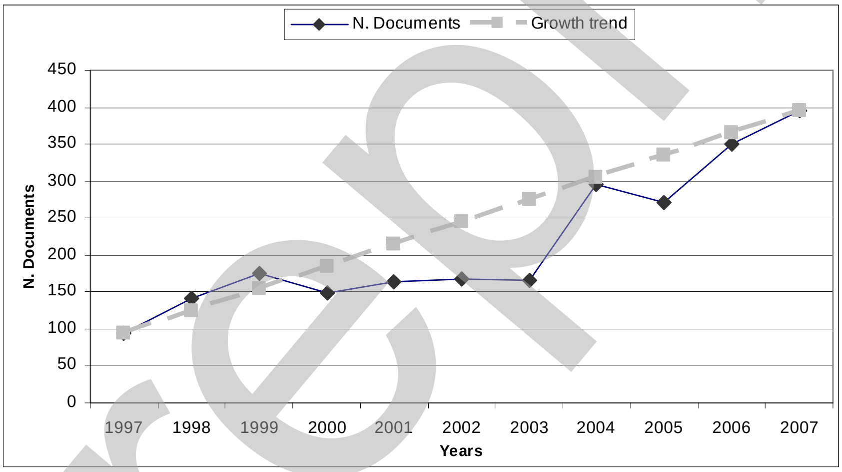

Figure 1. Evolution of output trend

5.1.2 Distribution by Autonomous Communities of Spain

Out of the seventeen regional "Autonomous Communities" (AC) into which Spain is divided, there are five that contributed most substantially to output, according to author affiliation (Figure 2); they are Cataluña, Madrid, Comunidad Valenciana, Castilla y León and Andalucía. The first two stand out way above the rest, with well over half of total production (67.2 %). Comunidad Valenciana has a share of 16%, while Castilla y León and Andalucía both contribute roughly 13% to total output, at somewhat of a distance from the rest of the AC.

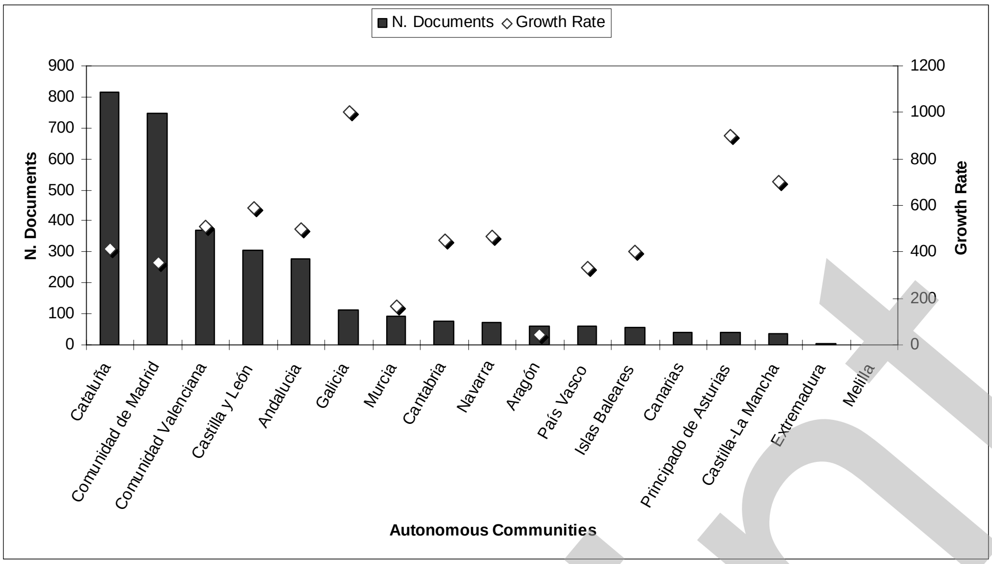

Figure 2. Distribution of output and growth rate by AC

Of the 644 institutions where national output originated, we see in Table 2 that the top 20 produce 50 publications in the period of study. Cataluña has two main sources of publications, the Hospital de Santa Cruz y San Pablo, with nearly 12% of output, and the Hospital Clínico de Barcelona and the Hospital Clínico de Salamanca, with over 8%. Following them, the most productive institution would be the Hospital la Fe of Valencia, with its roughly 5.5% contribution to output.

Institutions	N. Documents	% Documents	Growth Rate
Hosp. Sta. Cruz y San Pablo (Barcelona)	283	11.95	192
Hosp. Clin. De Barcelona.	197	8.32	900
Hosp. Clin. (Salamanca)	194	8.19	1000
Hosp. La Fe (Valencia)	130	5.49	767
IDIBAPS (Barcelona)	115	4.85	
Hosp. Univ. La Princesa (Madrid)	95	4.01	400
Hosp. Gregorio Marañón (Madrid)	87	3.67	
Hosp. La Paz (Madrid)	86	3.63	600
Hosp. Vall del Hebrón (Barcelona)	84	3.55	600
Hosp. Ramón y Cajal (Madrid)	77	3.25	600
Hosp. 12 Octubre (Madrid)	76	3.21	533
Hosp. Clin. (Valencia)	74	3.12	
Ctr. Inv. Energet. Medioamb. Tecnol. (Madrid)	68	2.87	100
Hosp. Marques de Valdecilla (Cantabria)	67	2.83	1000
Hosp. Niño Jesús (Madrid)	65	2.74	200
Inst. Catalán de Oncología (Barcelona)	57	2.41	0
Clin. Univ. Navarra (Navarra)	56	2.36	333
Ctr. Inv. del Cáncer (Salamanca)	54	2.28	350
Hosp. Reina Sofía (Córdoba)	54	2.28	
Hosp. Germans Trías Pujol (Barcelona)	51	2.15	1400

Table 2. Distribution of institutional output.

Also deserving mention is the output from the Instituto de Investigaciones Biomédicas August Pi i Sunyer of Barcelona (IDIBAPS). This "mixed institution" of administrative multi-dependence is dependent on the Hospital Clínico de Barcelona, Universidad de Barcelona, Inst. de Invest. Biomédicas de Barcelona del CSIC (IIBB-CSIC) and the Generalitat de Cataluña (Department of Economía y Conocimiento (DEC), and is physically located in the Hospital Clínico de Barcelona.

Table 3 shows how most of Spain´s research with stem cells is carried out in the Health Services. Approximately 64% of the publications can be traced to health institutions, which is twice as much as the output from the university sector, second in production. Other productive institutions are those that depend on the Administration of the Autonomous Communities (7.34%), the mixed institutions of the CSIC-Higher Education (6.37%) and the *Consejo Superior de Investigaciones Científicas* (CSIC) itself (5.61%). Documents from the Central Government or *Administración Central* (3.63%) correspond almost exclusively to the Instituto de Salud Carlos III, which at that time was governed under the Ministerio de Sanidad (Ministry of Health).

Institutional Sector	N. Documents	% Documents	Growth Rate
Health	1515	63.95	282
Higher Education	692	29.21	337
Public Administration	174	7.34	1700
CSIC-Higher Education	151	6.37	540
Consejo Superior de Investigaciones Científicas_CSIC	133	5.61	833
Multi-dependent Mixed Institutions	121	5.11	
Central Administration	86	3.63	600
Private	29	1.22	
CSIC-Higher Education (Local)	15	0.63	
State-Owned Companies	8	0.34	
Foundations	2	0.08	
Local Administration	1	0.04	
CSIC-Autonomous Communities	1	0.04	

Table 3. Distribution and growth of output by institutional sector

5.1.3 Most productive authors

Table 4 shows, in descending order, the percentage of documents that the authors participated in, along with the rate of growth over the period on the whole. The most productive authors, J. Sierra, J. F. San Miguel and R. Martino, have output that surpasses 120 documents.

Author	N. Documents	% Documents	Growth Rate
Sierra, J.	155	6.54	80
San Miguel, J. F.	143	6.04	1150
Martino, R.	140	5.91	67
Caballero, M. D.	126	5.32	33
Sureda, A.	120	5.07	137
Carreras, E.	112	4.73	150
Brunet, S.	100	4.22	-11
Montserrat, E.	98	4.14	120
Urbano-Ispizua, A.	94	3.97	150
Rovira, M.	83	3.50	125
Perez-Simon, J. A.	72	3.04	1000
De La Rubia, J.	64	2.70	1300
Ribera, J. M.	60	2.53	
Del Canizo, M. C.	58	2.45	600
Lahuerta, J. J.	56	2.36	250

Table 4. Most productive authors

The total sum of all authors signing documents comes to 18,571. This figure breaks down to a total of 7,397 actual individual authors, since many authors sign more than one paper. As some studies have reported (Newman 2001), this field presents a very high collaboration rate. The mean number of authors per paper in the biomedical fields focused in Medline database was 3.75 and concretely in cardiovascular subfield 6.18 (Bordons *et al.,* 1996). However, our data show that during the period of study (Figure 3), the mean authorship per document (co-authorship index) was higher with a rate of 7.8; this figure was quite stable over time (line with squares). This represents a strong pattern of coauthorship, even higher than the global of the biomedical subfields as Cronin pointed out at the beginning of 2000 (Cronin 2001), and this rate is similar to that found in the study of Zhao and Strotmann (2011).

Another indicator associated with productivity is the number of documents per author (dotted line with diamonds), which is seen to also be fairly stable over time, with a mean of five documents per author (dotted line with squares), though there is a slight increase in the middle of the study period. In contrast, the trend of authors per year (line with triangles) is on the rise, with a noteworthy surge in the year 2004.

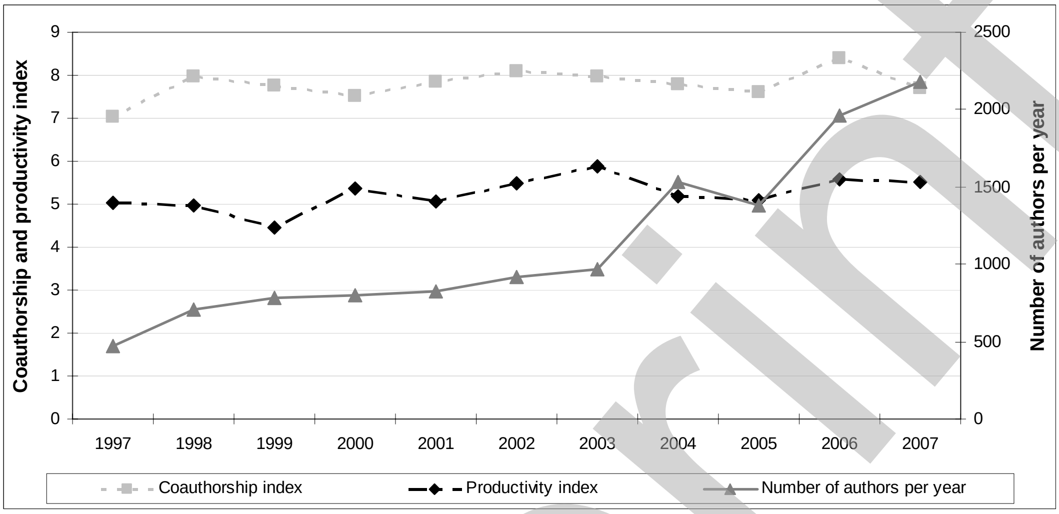

Figure 3. Index of co-authorship, productivity and evolution of the number of authors by year.

The co-authorship and productivity indicators reflect a coherent relation and a stable trend within the period of study, and trace a logical evolution towards an increase in co-authored documents at the expense of those put out by a single author. In fact, the transitory index, or authors with just one document, is 69%. In Figure 4 we find the distribution of the number of documents at each level of aggregation for the authors, Spanish institutions and foreign countries. At the micro level, the documents signed by 4, 5 or 6 authors constitute the bulk of output, which is coherent with the average number of authors per document that we saw in the previous figure. Therefore, we may affirm that the number of co-authored documents is considerably greater than the number of documents with a single author. Multiple authorship (more than two authors) is particularly prevalent, with a great proportion of documents signed by 11 to 20 authors.

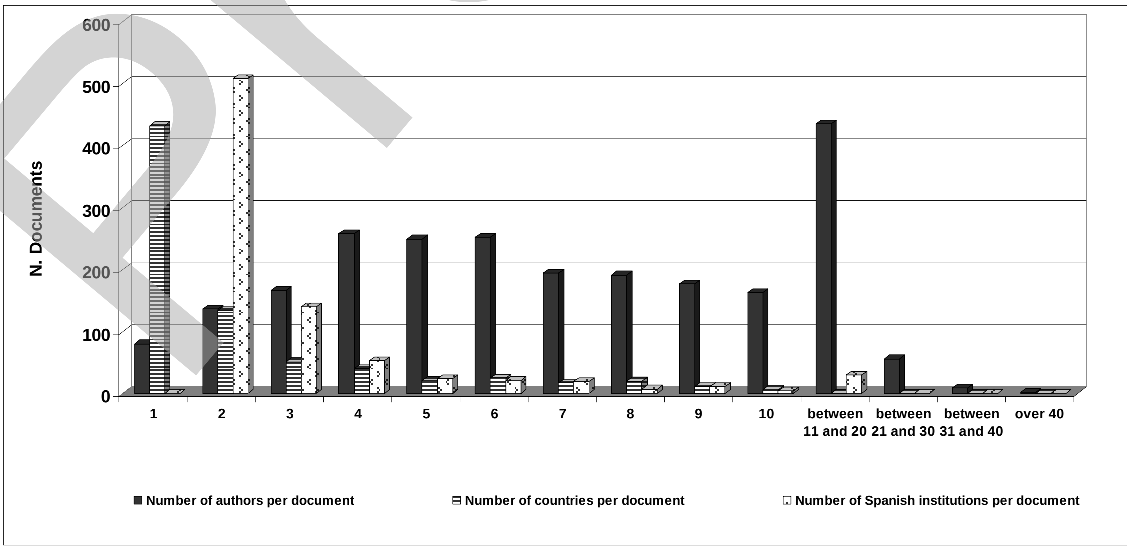

Figure 4. Percentage of documents signed in collaboration, according to number of participating actors

With respect to the number of institutions participating, we see that over 20% of output is signed by two institutions. Similarly, international collaboration amounts to roughly 18% of output signed by another country in conjunction with Spain. Trilateral or multilateral collaboration on the institutional or international level is very scarce. Therefore, when there is institutional and international collaboration, studies of a bilateral origin prevail over those of a trilateral or quadrilateral nature.

5.1.4 Patterns of collaboration

Analysis of the different types of collaboration (Figure 5) reveals patterns that tend to match those observed at the national or international level. As the period advances, the documents involving no collaboration descend, while the documents undersigned by more than one institution (national or international) are on the rise. These rates are higher than the ones reported by Zhao and Strotmann for the period 2004-2009. This might be due to differences between periods and sources. In any case, it would be desirable correct this tendency to avoid fluctuations over time and attain reasonable rates in accordance with international standards and patterns (less than 10%), bearing in mind the positive correlation between institutional and international collaboration and the high number of citations of collaborating authors as compared to single authorship (Narin, Stevens and Whitlow, 1991, Chinchilla *et al.,* 2010)

There is a remarkable tendency in the year 2004 for documents in collaboration with Spanish institutions to drop. Likewise, international collaboration decreases considerably the following year, 2005. Yet if we place these findings in the context of more authors per year, as reflected in Figure 3, one may infer greater interdepartmental collaboration is underway. Further study, at the level of institutional desegregation, would be necessary to confirm this inferred behaviour.

Therefore, if we relate these collaboration indicators with the overall output trend depicted in Figure 1, it is difficult to attribute the increase in publications detected in 2004 to such causes.

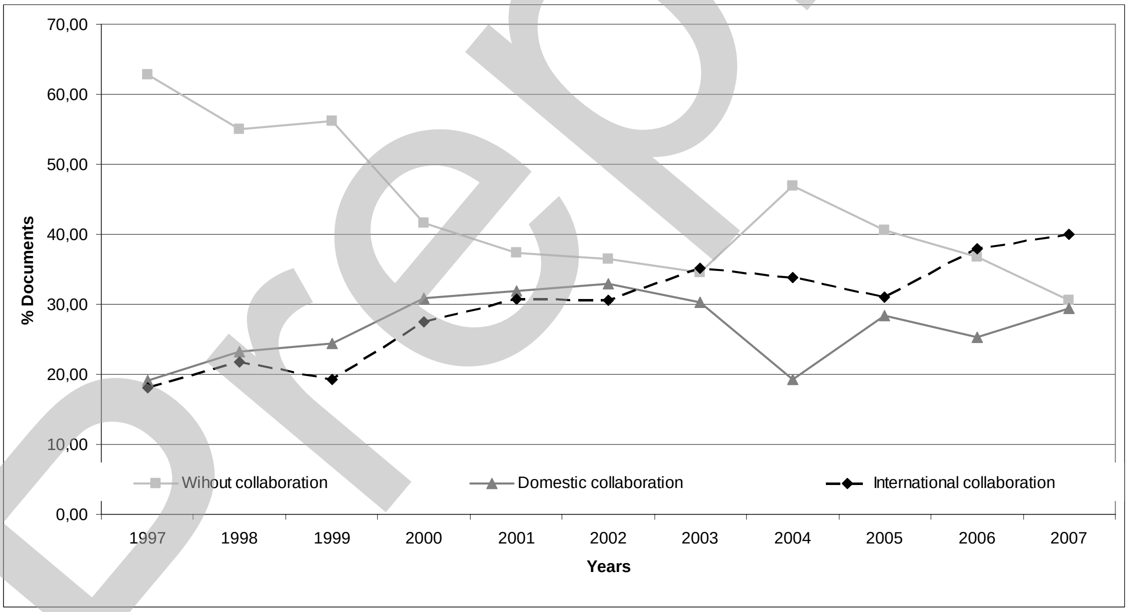

Figure 5. Types of collaboration

In the period overall, 31.95% of the documents analysed correspond to international collaborative efforts. The main countries participating are the United States, Germany, Great Britain, France, and Italy, followed by production involving modest but important countries in the area of stem cell research as can be seen in Table 5. There are noteworthy rates of growth in the two main countries of association, followed by Switzerland and Canada.

Country	N. Documents	% Documents	Growth Rate
USA	304	40.43	1220
Germany	183	24.34	1450

United Kingdom	178	23.67	900
France	174	23.14	420
Italy	156	20.74	833
Netherlands	109	14.49	800
Sweden	69	9.18	800
Switzerland	69	9.18	1000
Belgium	52	6.91	175
Canada	48	6.38	1200

Table 5. Main countries associated in Spain´s international collaboration and output in stem cell research

5.1.5 Document type

The predominating document type is the research article (1,555 papers), followed by congress proceedings, as seen in Figure 6. Deserving mention is the high number of reviews on the subject of stem cell research. In the period studied, 95% of output was in the form of a scientific article (over 65%), congress proceedings (20%) and reviews (8.82%). Over the years of study, however, there are some variations in the production by type of document: the growing number of congress proceedings and the reviews contribute to a relative decrease in the percentage of original research articles. An increase in the latter is seen over the period of study, especially alter 2004. Regardless, this finding can not be singled out as the cause of the fluctuations in total output observed in Figure 1, as the number of publications in the article format also adopts a growing trend.

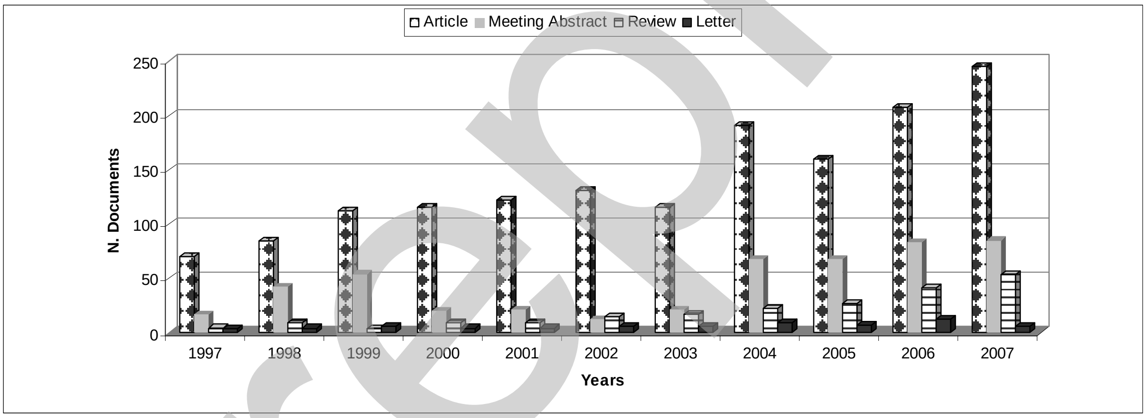

Figure 6. Document type

As the patterns of collaboration and the document type do not, in themselves, explain the increase in total output seen in the year 2004**,**_we also must consider that legislative reforms undertaken by the Spanish institutions could have implications for research, propelling it toward new and diversified fields, enriching the thematic profile of this scientific domain. This explanation is quite probable, as in recent years stem cell research has generated much debate in nearly all the sectors involved in biomedical research and development. For instance, at the legislative level, key reforms came about such as the Law of Assisted Reproduction on 21 November 2003, by virtue of which the "Commission for the Follow-up and Control of the Donation of Human Cells and Tissues" was created, which in turn authorized, in 2005, the very first research projects with embryonic stem cells. This law was derogated by the one in force, from 26 May 2006; among other things, it authorized the performance of autologous transplants from umbilical cord blood, allowing for the creation of private blood and cell banks. Another recent example is Law 14/2007 (3 July 2007) on Biomedical Research, which authorizes therapeutic cloning and research with embryos.

5.1.6 Journals of publication

The total number of journals where the documents were published came to 610. Outstanding among them is the output in *Bone Marrow Transplantation* and *Blood*, where practically 25% of the total appears. The rest are at a great distance in terms of the volume of production, and 97.7% of the journals show fewer than 20 publications involving stem cell research and Spain. Table 6 shows the journals that put out more than 10 documents during the study period, accounting for just over half of production.

Journal	N. Documents	% Documents	Growth Rate	Impact Factor 2006
Bone Marrow Transplantation	328	13.85	80	2.62
Blood	249	10.51	50	10.37
Haematologica	75	3.17	-100	
British Journal of Haematology	56	2.36	-33	4.50
Medicina Clinica	44	1.86	-50	1.33
Haematologica-the Hematology Journal	40	1.69	380	5.03
Leukemia	33	1.39	800	6.15
Annals of Oncology	31	1.31	0	5.18
Experimental Hematology	31	1.31	200	3.41
Biology of Blood and Marrow Transplantation	29	1.22	400	3.46
Journal of Neuroscience	26	1.10	200	7.45
European Journal of Cancer	24	1.01	-86	4.17
Journal of Clinical Oncology	21	0.89	100	13.60
European Journal of Neuroscience	20	0.84	0	3.71
Journal of Comparative Neurology	18	0.76	0	3.83
Leukemia & Lymphoma	18	0.76	100	1.56
Transfusion	18	0.76	-50	3.28
Brain Research	15	0.63	100	2.34
Histology and Histopathology	15	0.63	100	2.18
Revista De Neurologia	15	0.63	0	0.53
Proceedings of the National Academy of Sciences of the United States of America	13	0.55	200	9.64
Cytotherapy	12	0.51	300	2.16
Journal of Biological Chemistry	12	0.51	100	5.81
Stem Cells	12	0.51	200	7.92
Transplantation Proceedings	12	0.51	-33	0.96
Annals of Hematology	11	0.46	-67	2.25
Brain Research Bulletin	11	0.46	-100	1.68
European Journal of Haematology	11	0.46	200	1.86
International Journal of Developmental Biology	11	0.46	300	3.58
Physiologia Plantarum	11	0.46	-33	2.17
Revista Espanola De Cardiologia	11	0.46	200	2.18

Table 6. Most productive journals (more than 10 documents)

We add a column with the Impact Factor for the year 2006 to orient the reader, as it is impossible to compare journals on an equal basis, and so to highlight the amount of output in view of each journal´s visibility in terms of the IF. Thus we see that one of the most visible journals in the area surrounding stem cell research is second in production.

5.2 Thematic Analysis

5.2.1. Analysis of Subject categories

Altogether, the documents studied were published in journals classified under 110 different categories. The category *Hematology* stands out with its 997 documents, which account for 42% of the total. Next in line would be: *Oncology, Biophysics, Neurosciences, Cell Biology, Biochemistry and Molecular Biology* and *Immunology*. We would add that a major part of this production comes from the documents appearing in the journals *Bone Marrow Trasplantation* (*Hematology, Oncology and Biophysics*) and *Blood* (*Hematology*) where the majority of these documents are concentrated and as shown in Table 6

Subject categories	N. Documents	% Documents	Growth Rate
Hematology	997	42.09	209
Oncology	570	24.06	161
Biophysics	351	14.82	120
Neurosciences	215	9.08	367
Cell Biology	181	7.64	1175
Biochemistry & Molecular Biology	142	5.99	617
Immunology	117	4.94	2600
Biotechnology & Applied Microbiology	99	4.18	2300
Medicine, Research & Experimental	91	3.84	950
Plant Sciences	77	3.25	40
Transplantation	68	2.87	
Medicine, General & Internal	62	2.62	-50
Genetics & Heredity	60	2.53	
Clinical Neurology	59	2.49	
Developmental Biology	55	2.32	333
Pathology	52	2.20	433
Pharmacology & Pharmacy	49	2.07	200
Cardiac & Cardiovascular Systems	48	2.03	
Surgery	39	1.65	
Microbiology	37	1.56	100
Infectious Diseases	33	1.39	
Endocrinology & Metabolism	30	1.27	
Anatomy & Morphology	23	0.97	400
Multidisciplinary Sciences	23	0.97	500
Biochemical Research Methods	22	0.93	
Dermatology	21	0.89	200
Radiology, Nuclear Medicine & Medical Imaging	21	0.89	100
Zoology	21	0.89	0
Biology	18	0.76	300
Peripheral Vascular Disease	18	0.76	

Table 7. Most productive subject categories

A look at the evolution of these categories with regard to total productions reveals the most productive ones to be responsible for the overall trend in production (see Figure 1). *Neurosciences* constitutes a very productive category throughout the period of study; yet from 2004 onward, its output rises even more dramatically. Similarly, *Cell Biology*, *Biochemistry and Molecular Biology, Biotechnology and Applied Microbiology* and *Immunology* undergo sharp rises. In this comparative framework, certain parallels can be drawn with *Hematology, Oncology* and *Biophysics*. Thus, in Figure 8 we remove these three categories to emphasise the weight and scientific influence of the rest.

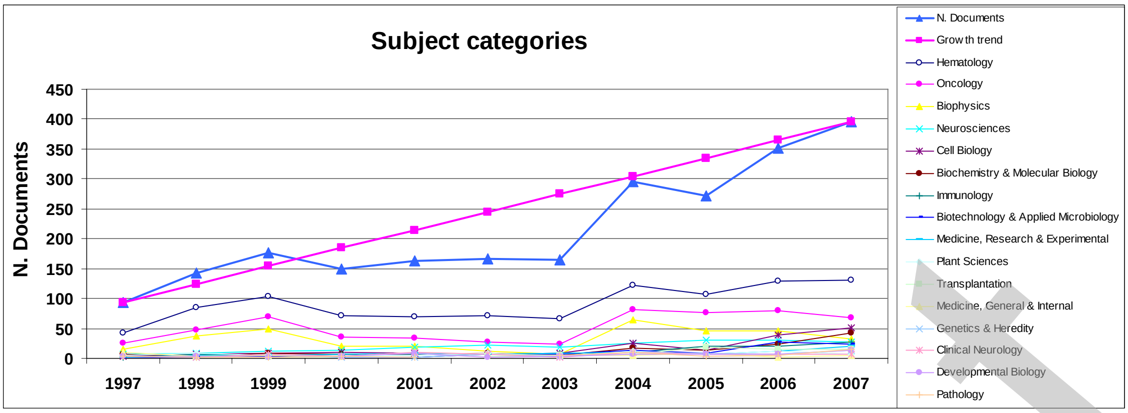

Figure 7. Evolution of the most frequent categories

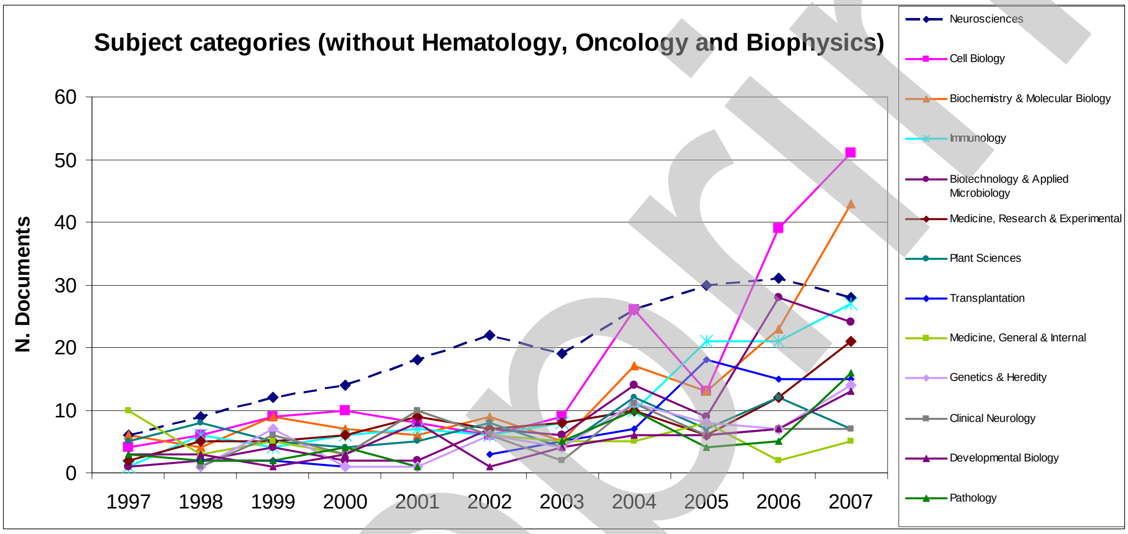

Figure 8. Evolution of the most frequent categories without Hematology, Oncology and Biophysics

Only the most productive ones are represented (over 50 documents).

Up to now, the most solid trend derived from our findings is that most output comes from Cataluña and Madrid, and is fundamentally produced in the health sector. More specifically, the outstanding centers would be: Hospital Sta. Cruz and S. Pablo de Barcelona, Hospital Clínico de Barcelona and the Hospital Clínico de Salamanca; and that the most productive authors belong to these centres, namely J. Sierra, and R. Martino, and J.F. San Miguel. We also corroborated that the most productive journals correspond precisely with the categories of greatest output. All this leads us to designate one single research area in Spain as the one mainly related with stem cell research —Hematology and Oncology. That is, studies involving clinical research are foremost in the field and in the literature. Its weight is so substantial that, when we carried out the bibliometric study using indicators, it was not possible to detect the implications of other less productive disciplines or research areas. We shall proceed to see if, through the representation of the relations established among the most productive categories, we might detect disciplines other than Oncology and Hematology involved in the bulk of stem cell research in Spain.

- ***Visual Subject Category Analysis***

Figure 9 displays the network of the 29 most frequent categories in the journals publishing the 2,369 documents retrieved with reference to stem cell research.

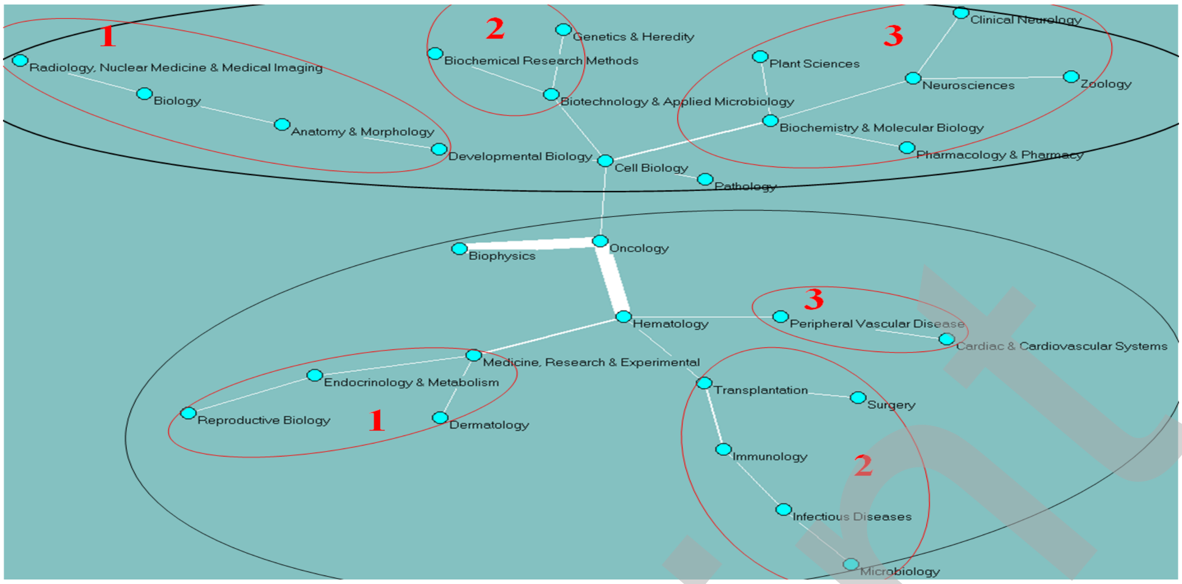

Figure 9. PFNET network of the most frequent categories related with stem cells.

The map of categories clearly shows a single main nucleus comprising Hematology, Oncology and Biophysics, where the most solid (red) relations are seen. Indeed, the most productive journal, Bone Marrow Transplantation (328 documents), is classified in these three categories.

Aside from this main nucleus, we see that Oncology gives rise to two rather weak groupings. The one toward the top of the map contains Cell Biology, which in turn embraces four subgroups, namely: Developmental Biology, Pathology, Biochemistry and Molecular Biology and Biotechnology and Applied Microbiology. In the lower part of the map, Hematology is a category that shows three other subgroupings, in this case: Medicine, Research and Experimentation, Peripheral Vascular Disease and Transplant.

Unlike what we have seen so far, this map allows us to distinguish two types of research, clinical and basic. The categories related with the former are situated in the central and lower part of the map and they correspond with the thematic groupings detected on the one hand, such as Hematology, Oncology and Biophysics, where the most solid (red) relations are seen and on the other, Medicine, Research and Experimentation, or Peripheral Vascular Disease and Transplant. The categories related with basic research, in contrast, are found in the upper area of the map, grouped under Cell Biology, which in turn harbors a number of subgroups.

However, this category analysis provides information more closely related with the subject classification of the journals where the documents are published (Moya-Anegón *et al.*, 2004) than with the information contained in the documents *per se*. We must bear in mind, then, that the data and conclusions to be deduced from this type of analysis serve only for journal characterization, not for any thematic breakdown. The type of information it offers is of a general character, and does not allow us to descend to a lower level of aggregation that might enable us to detect groupings with sufficient identity to reflect more specific areas or lines of research. Therefore, we shall use the KW+ as minor units of analysis that try to capture the intellectual content of documents so that one may visualize the subject relations existing among them.

5.2.2 KW+ analysis

Of the 2,369 documents obtained, 2,077 were indexed in the database using some sort of descriptor (1,846 documents had KW+ and/or AKW, and 231 had only the AKW). This means that 78% of the publications are indexed under KW+ descriptors —a remarkably high documental presence..The total number of KW+ was 16,024, of which 6,199 KW+ were unique. The mean number of the descriptors per document is 8.68 when considered in terms of the documents with some type of KeyWords (1,846).

	N. Documents	
Total	2,369	
With descriptors	2,077	
	With KW+	Only with AKW
	1,846	231

Table 8. Documents indexed

	N. KW+
Total	16,024
Unique	6,199
Nº of KW+ per document (n=1846)	8.68

Table 9. Nº of KW+

Descriptor	Frequency	% Documents with KW+*
Bone-Marrow-Transplantation	390	21.13
Stem-Cells	207	11.21
Expression	160	8.67
Stem-Cell Transplantation	142	7.69
Therapy	124	6.72
Progenitor Cells	123	6.66
Central-Nervous-System	113	6.12
Colony-Stimulating Factor	106	5.74
Bone-Marrow	103	5.58
Differentiation	103	5.58
Transplantation	99	5.36
In-Vitro	97	5.25
Graft-Versus-Host-Disease	94	5.09
Chemotherapy	80	4.33
Gene-Expression	78	4.23
Hematopoietic Stem-Cells	73	3.95
Leukemia	69	3.74
High-Dose Chemotherapy	66	3.58
Engraftment	65	3.52
Randomized-Trial	65	3.52
Peripheral-Blood	65	3.52
Survival	64	3.47
In-Vivo	61	3.30
Blood	60	3.25
Brain-Stem	59	3.20
Disease	58	3.14
Embryonic Stem-Cells	55	2.98
Mice	54	2.93
Recipients	53	2.87
Acute Myeloid-Leukemia	52	2.82
Non-Hodgkins-Lymphoma	51	2.76

Table 10. Most frequent descriptors. Note: *in view of docs that have KW+ (1.846)

Regarding the prevalence of appearance of the KW+, the most frequent descriptor is *Bone-Marrow-Transplantation* with 390 hits, nearly twice the yield of *stem cell,* with 207. Although the term used in our search strategy was "stem cell" in the field "topic", it retrieved documents that included this term in the KW Author, Title, Abstract, and of course, KW+ as well. For this reason, the term appearing the most times in the KW+ is not necessarily the search term. *Expression, Stem-Cell Transplantation, Therapy and Progenitor Cells* are found over 120 times, while *Central Nervous-System, Colony-Stimulating Factor, Bone-Marrow and Differentiation* close the group of descriptors with frequencies of appearance over 50 (Table 10).

➢ ***Co-word network***

Figure 10 offers a heat map where the main lines of research stand out. This visualization was obtained using VOSviewer, taking as the matrix of co-occurrence of the 6,199 unique KW+s, standardized by the measure of strength of association (Van Eck and Waltman, 2009).

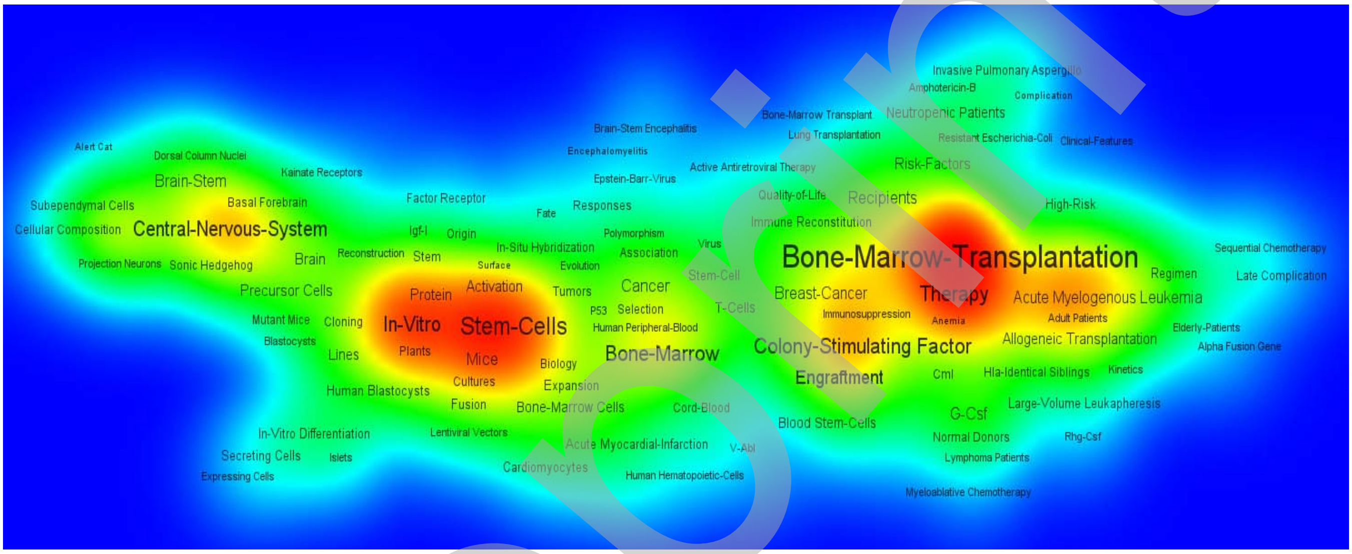

Figure 10. Visualization of the main lines of research involving stem cells (VOSviewer 1.3)

Although the KW+, which represent the domain of stem cells, show a highly interrelated distribution, the structure can be clearly seen. To the right of the map the documents related with the therapeutic use of hematopoietic progenitor stem cells are grouped, as indicated by the weight of *Therapy* in this KW+ set. As expected, this core appears linked to terms related to different clinical aspects of hematopoietic stem cell transplantation such as diseases (i.e. leukemia, lymphoma, breast cancer…), technical aspects (i.e. allogeneic, leukapheresis, immunosuppression…), outcome and complications. In the central position we see a concentration of documents referring to more basic research, reflected by the weight of the term *In-vitro* and animal experimental research (i.e. mice). Peripheral to this core we find terms related to early applications of stem cell research for regenerative medicine (i.e. cardiomyocytes and acute myocardial infarction) as well as the term cancer, presumably reflecting studies on cancer stem cell biology. Finally, to the left works involving stem cells in the context of neurosciences appear clustered and well defined.

As we stated earlier, the visualization obtained with VOSviewer permits quick and easy identification of the main areas of research in this domain. However, if we also wish to view the most relevant descriptions within each line of research, we will need to resort to other types of visualization, based on alternative techniques and methods.

Figure 11 shows the PFNET network of the 102 KW+ that appear most frequently in the 2,369 documents retrieved for this study. The representativeness of these descriptors is quite high, as they are present in 68.80% of the documents. This finding confirms the validity and adequacy of the method in locating and representing major areas of research, despite the fact that the degree of dispersion of the KW+ can be greater than that of other indexing systems.

Analysis of the relations among the most frequent descriptors reveals four well defined groups, but with a variable degree of dispersion. In the upper left part, around the node *Bone Marrow*

Transplantation, we find the most compact group of the network. It includes descriptors that represent documents in the areas of Hematology and Oncology, related fundamentally with the application and clinical research into the transplant of hematopoietic progenitor cells begun in the 1980's. From the descriptor *Colony-Stimulating Factor*, three other groupings appear. To the left, the KW+ *stem cells* connects with a well defined cluster of descriptors related with hematopoietic stem cell research (*Bone Marrow, Placental Blood, Hematopoietic stem cells, Transplantation, Cancer*…) Meanwhile, the term *Progenitor Cells* branches out into two other lines. To the right, we have the KW+ associated with basic research related with the processes of generation, proliferation and the differentiation of lines of stem cells (*Embryonic stem-cells, Mesenchymal stem-cells, Gene-Expression, Cell, Protein, Growth, Culture, Lines, Self-Renewal, Proliferation, Activation, Apoptosis*) Finally, in the lower zone and around the descriptor *Central-Nervous-System* we can clearly discern a line grouping descriptors specifically related with neural progenitor cells (*Neural Stem-Cells*, *Neurogenesis, Spinal-Cord, Brain-Stem, Neurons…*), which reveals the special attention that they are given from the realm of the neurosciences.

Even though the depictions of Figures 10 and 11 were made using different programs, data, techniques, methods and layout algorithms, the two coincide in identifying the number and name of the main areas of research surrounding stem cells in Spain.

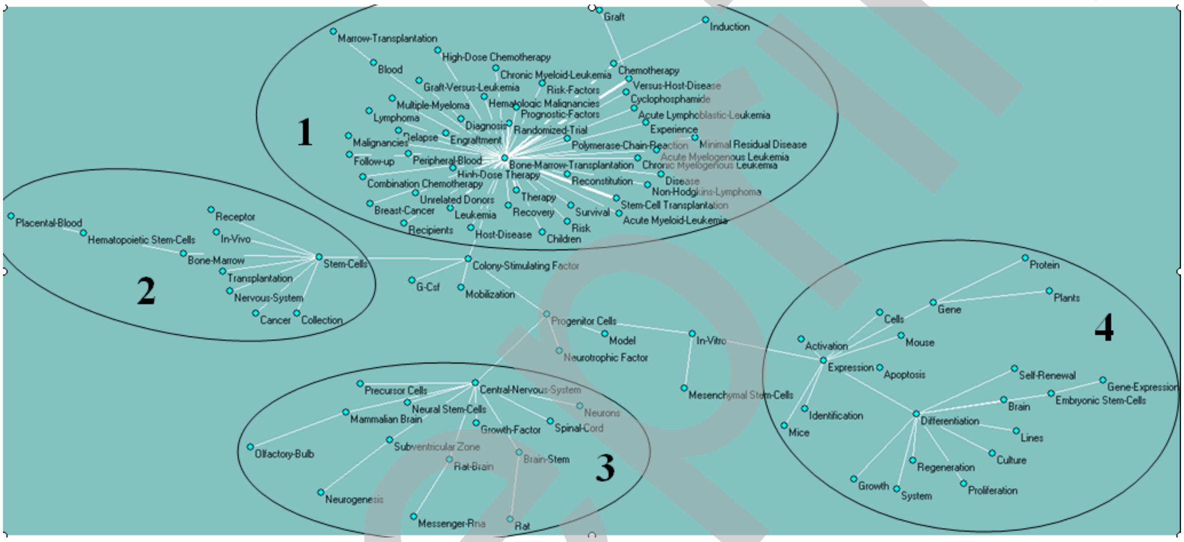

Figure 11. PFNET of the 102 most frequent KW+

In addition, the results obtained from category analysis in the previous section coincide with those derived from the KW+, indicating that the most weight in stem cell research in Spain resides in hematological and oncological studies. This is an interesting finding because, despite the theoretically highly interdisciplinary nature of the field, this area is reportedly dominated by a few central medical research areas (Zhao and Strotmann, 2011). Category analysis does not allow us to define research lines beyond the set of disciplines directly involved in the domain of study. For this reason, the information obtained through KW+ analysis is much more comprehensive and precise. Even though the two largely coincide or overlap, that of subject categories (Figure 9) is more general or superficial than the information based on KW+ (Figures 10 and 11). The latter show, with a greater level of desegregation, the distribution of the descriptors that specifically configure each thematic profile, plus the less productive or incipient research that would otherwise remain hidden.

6. Conclusions and Further Research

The results we harvested and present here give substantial information about the state of research involving stem cells in Spain to date, and allow us to draw profiles from different viewpoints.

Firstly, bibliometric analysis, which relies on a general search strategy, reflects a series of data that serve to characterize the scientific domain and lend it an identity of its own to better contextualize the study. This provides the reader or the decision-maker with a point of departure for grasping the domain. This study has allowed us to depict a general panorama of Spain´s research efforts surrounding stem cells by means of a bibliometric analysis, and evidences the possible influence of certain patterns of publication and collaboration. It also makes it possible to detect and represent the research trends and areas that characterize this particular scientific domain.

As could be expected, the thematic analysis based on the study of categories offers much more general, superficial and imprecise information than the KW+. The choice of categories as the units of measure for representing thematic relations among documents is very promising indeed, as it can have a dual function. It can be used to locate the main lines of research under a given scientific domain, and can moreover serve as a means to locate the documents pertaining to specific research lines, allowing for their direct application to carry out bibliometric studies at the micro level.

We are, of course, aware of the fact that this methodology has its limitations, which should be brought to light here for the benefit of future work. Granted, using one or another data source may condition the results of the study. Database coverage is determinant for obtaining more or less comprehensive results**.** Future studies may therefore lead us to apply this same methodology to alternative databases, so as to arrive at comparative results...

The use of different information sources world furthermore serve to analyze the indexing systems from diverse perspectives. Firstly, using the same systems and the same documental or disciplinary set (for instance, AKW) could lead us to uncover differences in data representation. Secondly, using different indexing systems might help us perceive the different foci of interest and the emergence or consolidation of research lines, depending on the use of controlled or non-controlled vocabularies.

We should moreover mention that the search strategy applied here was not chosen for the purpose of retrieving all the Spanish research papers on stem cells publisher, but rather to retrieve a volume of documents that would be heterogeneous enough to serve as a pool of studies from which to discern research lines through thematic analysis. In the near future, it would indeed prove interesting to take the most frequent KW+ from each research line detected, and then relaunch them onto the database. In this way, the retrieval of data on this specific subject would be more exhaustive, and the results could be compared with those obtained using other search (categories, keywords, topics…). Similarly, we could focus on the study of one or more lines detected, in order to develop bibliometric studies at the micro level.

Finally, we should acknowledge that the methodology applied and described here may entail limitations that could condition the results expounded. There are other analytical models (aside from co-words) that could reflect the research lines detected and offer additional, complementary information. Alternative or future techniques for visualization could likewise enhance the display of results.

In short, the results put forth here can be viewed as a composite still-life photograph of the period of study within the diffuse intellectual framework of Spain. It would prove very interesting to apply this methodology over a series of time periods to view the evolution of specific specialized areas and the degree of participation of the institutions, authors and geographic regions involved in biomedical research. The very nature of scientific collaboration makes it desirable to further explore these networks of co-authorship and author citation. We therefore stress that the notion of chronological studies incorporating new analytical and display techniques are a source of motivation for our research group.

For the time being, we conclude that the proposed methodology may prove very helpful for the thematic delimitation of stem cell research or any other scientific domain, and it can also be used as a tentative approach for the conceptual dissection of a domain, making manifest the main foci of interest.

References

Aharony, N. (2010), “Information literacy in the professional literature: an exploratory analysis”, *Aslib Proceedings*, Vol. 62 No. 3, pp. 261-282

Batagelj, V. and Mrvar, A. (2010), “Pajek 2.0: Package for large network analysis”, available http://vlado.fmf.uni-lj.si/pub/networks/pajek/ (accessed 8 March 2010).

Bellavista, J., Guardiola, E., Méndez, A. and Bordons, M (1997), *Evaluación de la investigación*, Centro de Investigaciones Sociológicas (Cuadernos Metodológicos), Madrid.

Bessalar, P., Heimeriks, G. (2006), "Mapping research topics using word-reference co-occurrences: A method and an exploratory case study", *Scientometrics*, Vol. 68 No. 3, pp. 377-393.

Bishop, AE., Buttery, L.D.K. and Polak J.M. (2002), "Embryonic stem cells", *Journal of Pathology*, Vol. 197 No. 4, pp. 424-429.

Bordons, M., Gómez, I., Fernández, M. T., Zulueta, M. A., Mendez, A. (1996). "Local, domestic and International scientific collaboration in biomedical research", *Scientometrics,* Vol. 37 No. 2, pp. 279-295.

Bordons, M. and Gómez Caridad, I. (1997), "La actividad científica española a través de indicadores bibliométricos en el período 1990-93", *Revista General de Información y Documentación*, Vol. 7, No. 2, pp. 69-86.

Bordons, M. and Zulueta, M.A. (1999), "Evaluation of the scientific activity through bibliometric indices", *Revista Española Cardiología*, Vol. 52 No. 10, pp. 780-800.

Börner, K., Chen, C., and Boyack, K.W. (2003), "Visualizing knowledge domains", *Annual Review of Information Science & Technology*, Vol. 37, pp. 179-255.

Boyack, K.W., Börner, K. and Klavans, R. (2009), "Mapping the structure and evolution of chemistry research", *Scientometrics*, Vol. 79, No. 1, pp. 45-60.

Boyack K.W. and Klavans R. (2010), "Co-Citation Analysis, Bibliographic Coupling, and Direct Citation: Which Citation Approach Represents the Research Front Most Accurately?", *Journal of the American Society for Information Science and Technology*, Vol. 61, No. 12, pp. 2389-2404.

Braun, T., Glänzel, W., and Schubert, A., How Balanced Is the Science Citation Index's Journal Coverage? – A Preliminary Overview of Macrolevel Statistical Data. Cronin, B. and Atkins, H. B. Eds. The Web of Knowledge: A Festschrift in Honor of Eugene Garfield; 2000; pp. 251-277. (Asist Monograph Series).

Cambrosio, A., Limoges, C., Courtial, J.P. and Laville, F. (1993), "Historial scientometrics? Mapping over 70 years of biological safety research with co-word analysis" *Scientometrics*, Vol. 27 No. 2, pp.119-143.

Camí, J., Zulueta, M.A., Fernández, M.I., Bordons, M. and Gómez, I. (1997), "Spanish scientific production in biomedicine and health sciences during the period 1990-1993 (Science Citation Index and Social Science Citation Index) and comparison to period 1986-1989" *Medicina Clinica*, Vol.109. No. 13, pp. 481-496.

Cao, Q.L, Benton, R.L. and Whittemore, S.R. (2002), "Stem cell repair of central nervous system injury", *Journal of Neuroscience Research*, Vol. 68, No. 5, pp. 501-510.

Chalik, T. and Jirina, M. (2006), "Law of cumulative advantages in the evolution of scientific fields", *Scientometrics*, Vol. 66 No. 3, pp. 441-449.

Chen, C., Ibekwe-SanJuan, F., and Hou, J. (2010), "The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis", *Journal of the American Society for Information Science and Technology*, Vol. 61, No. 7, pp. 1386-1409.

Chen D.Z., Huang M.H., Hsieh H.C., and Lin C.P., (2011), "Identifying missing relevant patent citation links by using bibliographic coupling in LED illuminating technology", *Journal of Informetrics*, Vol. 5, No. 3, pp. 400-412.

Chiau, W.T. and Ho, Y.S. (2007), "Bibliometric analysis of tsunami research", *Scientometrics*, Vol. 73, No. 1, pp. 3-17.

Chinchilla-Rodríguez, Z. *Análisis del dominio científico español: 1995-2002 (ISI, Web of Science).* PhD Dissertation, Universidad de Granada, 2005.

Chinchilla-Rodríguez, Z. and Moya-Anegón, F. (2007*), La investigación científica española (1995-2002): una aproximación métrica,* Universidad de Granada, Granada.

Chinchilla-Rodríguez Z, Vargas-Quesada B, Hassan-Montero Y, González-Molina A, and Moya-Anegón F. (2010), "New approach to the visualization of international scientific collaboration" *Information Visualization*, Vol. 9 No. 4, pp. 277-287.

Cobo M.J., López-Herrera A.G., Herrera-Viedma E., and Herrera F. (2011), "An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field", *Journal of Infometrics*, Vol. 5, No. 1, 146-166.

Courtial, P., Cahlik, T. and Callon, M. (1994), "A model for social interaction between cognition and action through a key-word simulation of knowledge growth" *Scientometrics*, Vol. 3, No. 2, pp. 173-192.

Cronin, B. (2001), "Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices", *Journal of the American Society for Information Science and Technology*, Vol. 52, No. 7, pp. 558-569.

Di Giorgio, F.P., Carrasco, M.A., Siao, M.C., Maniatis, and T. Eggan, K. (2007), "Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model" *Nature Neuroscience*, Vol. 10, No. 5, pp. 608-614.

Ding, Y., Chowdhurry, G. and Foo, S. (2001), "Bibliometric cartography of information retrieval research by using co-word analysis" *Information Processing & Management*, Vol. 37, No. 6, pp. 67-78.

España. Ley 45/2003, de 21 de noviembre, por la que se modifica la Ley 35/1988, de 22 de noviembre, sobre Técnicas de Reproducción Asistida. Boletín Oficial del Estado, 22 noviembre de 2003, núm. 280, 41.458-41.463. [Law on Techniques of Assisted Reproduction].

España. Ley 14/2006, de 26 de mayo, sobre técnicas de reproducción humana asistida. Boletín Oficial del Estado, 27 de mayo de 2006, núm. 126, 19947-19956. [Law on Techniques of Assisted Reproduction].

España. Ley 14/2007, de 3 de julio, de Investigación biomédica. Boletín Oficial del Estado, 4 de julio de 2007, núm. 159, 28.826-28.848. [Law on Biomedical Research].

Fernández, M.T., Cabrero, A., Zulueta, M.A. and Gómez I. (1993), Constructing a relational database for bibliometric analysis. *Research Evaluation*, 3(1), 55-62.

García-Guinea, J. and Ruis, J. D. (1998), "The Consequences of Publishing in Journals Written in Spanish in Spain", *Interciencia*, Vol. 23, No. 3, pp. 185-187.

Garfield, E. (1990), "Keywords plus-ISI´s breakthrough retrieval method. Part 1. Expanding your searching power on Current Contents on Diskette", *Current Contents*, Vol. l, No. 32, pp. 5-9.

Garfield, E. and Sher, I.H. (1993), "Keywords PlusTM – Algorithmic derivative indexing", *Journal of the American Society for Information Science*, Vol. 44, No. 5, pp. 298-299.

Glänzel W., Thijs B., (2011), "Using 'core documents' for the representation of clusters and topics", *Scientometrics*, Vol. 88, No. 1, pp. 297-309.

Hammond, J. (1949), "Recovery and culture of tubal mouse ova", *Nature*, Vol. 163, No. 4131, pp. 28-29.

Hjørland, B. and Albrechtsen, H. (1995), "Toward a new horizon in information science: domain analysis", *Journal of the American Society Information Science*, Vol. 46, pp. 400-425.

Ho, Y.S., Chiu, C.H., Tseng, T.M. and Chiu, W.T. (2003), “Assessing stem cell research productivity”, *Scientometrics,* Vol. 57, No.3, pp. 369-376.

Ho, Y.S. (2007), “Bibliometric analysis of adsorption technology in environmental science. *Journal of Environmental Protection Science”* Vol. 1, No. 1, pp. 1-11.

Jiménez-Contreras, E., Faba-Pérez C. and Moya-Anegón F. (2001), “El destino de la revistas científicas nacionales. El caso español a través de una muestra (1950-90)”, *Revista Española de Documentación Científica*, Vol. 24, No 2, pp. 147-161.

Jiménez Contreras, E., Moya Anegón, F., and Delgado López-Cózar, E. (2003), “The Evolution of Research Activity in Spain. The impact of the National Commission for the Evaluation of Research Activity (CNEAI)”, *Research Policy*. Vol. 32, No. 1, pp. 123-142.

Kamada, T. and Kawai, S. (1989), “An algorithm for drawing general undirected graphs”, *Information Processing Letters*, Vol. 31, No. 1, pp. 7-15.

Kessler, M. M. (1963), "Bibliographic coupling between scientific papers." *American Documentation*, Vol. 14, No. 1, pp. 10-25.

Krauskopf, D. (2000), “La construcción social de la Ciencia y la Tecnología”, *Boletín SEBBM*, Vol. 130, pp.12-16.

Kyvik, S. (2003), “Changing Trends in Publishing Behaviour among University Faculty, 1980-2000”, *Scientometrics*, Vol. 58, No. 1, pp. 35-48.

Lee, W.H. (2008), “How to identify emerging research fields using scientometrics: An example in the field of Information Security”, *Scientometrics*, Vol. 76, No. 3, pp. 503-525.

Leydesdorff, L. and Hellsten, I. (2005), “Metaphors and diaphors in science communication”, *Science Communication*, Vol. 27, No. 1, pp. 64-99.

Leydesdorff, L. and Hellsten, I. (2006), “Measuring the meaning of words in contexts: an automated analysis of controversies about “Monarch butterflies”, “Frankenfoods” and “stem cells””, *Scientometrics*, Vol. 67, No. 2, pp. 231–258.

Leydesdorff, L. and Rafols, I. (2009), “A global map of science based on the ISI subject categories”, *Journal of the American Society for Information Science and Technology*, Vol. 60, No. 2, pp. 348-362.

Leydesdorff, L. and Welbers, K. (2011), “The semantic mapping of words and co-words in contexts”, *Journal of Informetrics,* Vol. 5, No. 3, pp. 469-475.

Li, L.L., Ding, G., Feng, N., Wang, M.H. and Ho, Y.S. (2009), “Global stem cell research trend: Bibliometric analysis as a tool for mapping of trends from 1991 to 2006”. *Scientometrics*, Vol. 80, No.1, pp. 39-58.

Martínez Serrano, A. and Bjorklund, A. (1996), “Protection of the neostriatum against excitotoxic damage by neurotrophin-producing, genetically modified neural stem cells”, *Journal of Neuroscience*, Vol. 16, No. 15, pp. 4604-4616.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solan,a V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E. and Muñoz-Fernández, F.J. (2004), “A new technique for building maps of large scientific domains based on the cocitation of classes and categories”, *Scientometrics,* Vol. 61, No.1, pp. 29-145.

Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., González-Molina, A., Muñoz-Fernández, F.J. and Herrero-Solana, V. (2006), “Visualizing and analyzing the Spanish science structure: ISI Web of Science 1990-2005”, *Profesional de la Información*, Vol. 15, No. 4, pp. 258-269.

Moya-Anegón, F., Chinchilla Rodríguez, Z., Corera-Álvarez, E., Gómez-Crisóstomo, R., González-Molina, A., Muñoz-Fernández, F. and Vargas-Quesada, B. (2007), *Indicadores bibliométricos de la actividad científica española: 1990–2004*, Fundación Española para la Ciencia y la Tecnología, Madrid.

Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Munoz-Fernández, F.J. and Herrero-Solana, V. (2007), "Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology"*, Vol. 58, No. 14, pp. 2167–2179.

Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Gómez-Crisóstomo, R, Hassan-Montero, Y. and Vargas-Quesada, B. (2009), *Indicadores Bibliométricos de la Actividad Científica Española: 2007,* Fundación Española para la Ciencia y la Tecnología (Fecyt), Madrid.

Neff, M.W. and Corley, E.A. (2009), "35 years and 160.000 articles: A bibliometric exploration of the evolution of ecology" *Scientometrics*, Vol. 80, No. 3, pp.657-682.

Narin, F., Stevens, K., and Whitlow, E. S. (1991), "Scientific cooperation in Europe and the citation of multinationally authored papers", *Scientometrics,* Vol. 21, No. 3, pp.313-323.

Newman, M. E. J. (2001), "The structure of scientific collaboration networks" *Proceedings of the National Academy of Science of the USA*, Vol. 98, No. 2, pp.404-409.

Noyons, E. C. M., Moed, H. F. and Luwel M. (1999), "Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study", *Journal of the American Society for Information Science,* Vol. 50, No.2, pp. 115-131.

Onyancha, O.B. and Ocholla, D.N (2005), "Informetric investigation of the relatedness of opportunistic infections to HIV/AIDS" *Information Processing & Management*, Vol. 41, No. 6, pp.1573-1588.

Perianes-Rodríguez, A., Olmeda-Gómez, C., Ovalle-Perandones, M. A., Chinchilla-Rodríguez, Z., and Moya-Anegón, F. (2011), "R&D collaboration in 50 major Spanish companies", Aslib Proceedings, Vol. 63, No. 1, pp. 5-27.

Qin, J. (2000), "Semantic similarities between a keyword database and a controlled vocabulary database: An investigation in the antibiotic resistance literature", *Journal of the American Society for Information Science, Vol.* 51, No. 2, pp.166-180.

Quirin, A., Cordón, O., Guerrero-Bote, V.P., Vargas-Quesada, B. and Moya-Anegón, F. (2008), "A quick MST-based algorithm to obtain pathfinder networks (∞, n −1)", *Journal of the American Society for Information Science and Technology*, Vol. 59, No.12, pp.1912-1924.

Rafols, I., Porter, A. and Leydesdorff, L. (2010), "Science overlay maps: a new tool for research policy and library management", *Journal of the American Society for Information Science and Technology*, Vol. 61, No. 9, pp. 1871-1887.

Resolución de 28 de agosto de 1989, modificada y completada por el Real Decreto 1325/2002. . [Law on Performance Research].

Sanz-Casado, E., Aragon, I. and Mendez, A. (1995), "The Function of National Journals in Disseminating Applied Science", *Journal of Information Science*, Vol. 21, No. 4, pp. 319-323.

Small, H., and Upham, S. P. (2009), "Citation structure of an emerging research area on the verge of application", *Scientometrics*, Vol. 79, No. 2, pp. 365–375.

Su, H.N. and Lee, P.C. (2010), "Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in Technology Foresight", *Scientometrics*, Vol. 85, No. 1, pp. 65-79.

Thomson Reuters ISI Web of Science (2010). [Database online], available at: http://science.thomsonreuters.com/es/productos/wos (accessed 11 April 2008).

Van Eck, N.J. and Waltman, L. (2009), “How to normalize cooccurrence data? An analysis of some well-known similarity measures”, *Journal of the American Society for Information Science and Technology*, Vol. 60, No. 8, pp. 1635–1651.

Van Eck, N.J. and Waltman, L. (2010), “Software survey: VOSviewer, a computer program for bibliometric mapping”, *Scientometrics*, Vol 84, No. 2, pp. 523-538.

Van Raan, A.F.J. and Tijssen, R.J.W. (1993), “The neural net of neural network research. An exercise in bibliometric mapping”, *Scientometrics*, Vol. 26, No. 1, pp. 169-192.

Vargas-Quesada, B., Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E. and Guerrero-Bote, V. (2008), “Development of the Spanish scientific landscape: ISI web of science 1990-2005”. *Profesional de la Información*, Vol. 17, No. 1, pp. 22-37.

Vargas-Quesada, B., Chinchilla-Rodríguez, Z., González-Molina, A. and Moya-Anegón, F. (2010), “Showing the Essential Science Structure of a Scientific Domain and its Evolution”, *Information Visualization*, Vol. 9, pp. 288-300.

Wasserman, S. and Faust, K. (1998), *Social network analysis: methods and applications*, Cambridge University Press, Cambridge.

White, H. D and Griffith B. C. (1981), "Author co-citation: a literature measure of intellectual structure." *Journal of the American Society for Information Science*, Vol. 32, pp. 163-171.

White, H. and McCain, K. (1998), “Visualizing a Discipline: an Author Co-Citation Analysis of Information Science, 1972-1995”, *Journal of the American Society for Information Science,* Vol. 49, No.4, pp. 327-375.

White, H.D. (2003), “Pathfinder networks and author co-citation analysis: a remapping of paradigmatic information scientists”, *Journal of the American Society for Information Science and Technology*, Vol. 54, No. 5, pp. 423-434.

Zhao, D. and Strotmann, A. (2011), “Intellectual structure of stem cell research: a comprehensive author co-citation analysis of a highly collaborative and multidisciplinary field”, *Scientometrics,,* Vol.87, No.1, pp. 115-131.

Zitt, M., and Bassecoulard, E. (1996), “Reassessment of co-citation methods for science indicators: Effect of methods improving recall rates, *Scientometrics*, Vol. 37, No. 2, pp. 223–244.

Zitt, M., and Bassecoulard, E. (2006), “Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences”, *Information Processing & Management,* Vol. 42, No.6, pp. 1513–1531.

Zitt, M., Lelu, A., and Bassecoulard, E. (2008), “Hybrid maps of scientific fields (terms and citations): An application to nanosciences”, paper presented at the Tenth International Conference on Science&Technology Indicators: Excellence and Emergence—A New Challenge for the Combination of Quantitative and Qualitative Approaches, Vienna, Austria.

Zitt, M., Lelu, A., and Bassecoulard, E. (2011), “Hybrid citation-word representations in Science mapping: portolan charts or research fields?”, J*ournal of the American society for information Science and Technology,* Vol. 62, No 1, pp. 19-39.

Zulueta, M.A., Cantos-Mateos, G., Sánchez, C. and Vargas-Quesada, B. (2011), “Research involving women and health in the Medline database, 1965-2005, Co-term analysis and visualization of main lines of research”, *Scientometrics,* Vol. 88, No.3, pp. 679-706.