Bibliometric mapping as a science policy and research management tool

Noyons, E.C.M.

Author: Noyons, Ed C.M.
Title: Bibliometric mapping as a science policy and research management tool
Issue Date: 1999-12-09

4 Exploring the Science and Technology Interface: Inventor-Author Relations in Laser Medicine Research*

E.C.M. Noyons[a], A.F.J. van Raan[a], H. Grupp[b], and U. Schmoch[b]

[a] Centre for Science and Technology Studies (CWTS)
University of Leiden
Wassenaarseweg 52
P.O.Box 9555
2300 RB Leiden, The Netherlands

[b] Fraunhofer Institute for Systems and Innovation Research (FhG-ISI)
Breslauer Strasse 48
D-7500 Karlsruhe, Germany

---

* This research project was supported by two related research grants to CWTS and FhG-ISI from Volkswagen Foundation, Hanover. Reprinted from Research Policy 23, pp. 443-457, Copyright (1994), with permission of Elsevier Science.

Exploring the Science and Technology Interface: Inventor-Author Relations in Laser Medicine Research

Abstract

The aim of this study is to investigate a specific aspect of the science and technology interface: inventor-author relations. The subject area is application of lasers in medicine. The empirical material consists of a set of 30 patents, representing the 'technology side', and 1057 publications authored by the inventors, representing the 'science side' of lasers in medicine.

Our study includes four different approaches. First, we tried to find evidence, by looking at the scientific part, for the claim that references in patents to non-patent literature (NPL references, mostly scientific publications) indicate 'science intensity'. It appeared that inventors of patents with many NPL references did not publish significantly more in science than inventors of patents with few NPL references. The former did, however, use more basic scientific journals to publish in than the latter.

Second, we tried to identify at the science side one paper per patent which would best represent the R&D activities related to the patent. Here, a weak correlation was found between the number of NPL references in the patents and the number of references in their scientific counterparts.

In our third approach, we compared the number of NPL references in the patents with expert assessments about the science intensity of each individual patent. Moreover, other aspects were taken in consideration, such as legal status of a patent (number of claims), complexity of the invention (number of pages), size of the inventor team. We found out that some of these other aspects could be related to a higher number of NPL references in patents.

In the fourth and final approach of the study, we analyzed the inventors' publications in more detail, in particular for the period before and around the patent application date. We tested and found evidence for two hypotheses. These two hypotheses state that, in preparation of a patent application, (1) co-inventors increase their co-activity in science; and (2) companies and universities level up their co-operation.

4.1 Introduction

Science and technology constitute a complicated, heterogeneous system of activities characterized by many interrelated aspects [13]. Although, in principle, virtually all scientific (S) and all technological (T) activities might be connected in one way or another, and thus both domains of human knowledge form indeed one complex system, it is not unrealistic to hypothesize that both domains still have their own identity, characterized, however, by a specific interface region they have in common (Grupp [4]). This S&T interface is the nursery of research and development (R&D) activities, and it is undoubtedly a major driving force for the economic development of our world. Therefore, systematic investigation of this network of interrelations is of crucial importance for future R&D policy.

As research on the S&T interface is still in its exploratory stage, it is not surprising that there is hardly any thoroughly formalized and quantified evidence of 'science involvement' in innovation. The few exceptions are related to patent application examination at patent offices, in particular the search for 'prior art'. Often this 'prior art' involves earlier patent documents. However, occasionally, patent examiners document the earlier status of the subject area of invention by making references to scientific literature, and not to patent literature. If the number of such *non-patent literature* (NPL) references in patent documents is used as a measure of 'science involvement', a set of indicators can be constructed for quantifying this science involvement (Grupp and Schmoch [5,6]). Carpenter, Cooper and Narin [1] showed that US patents originating from specific 'scientific' areas of technology, contain significantly more references to the prior art which are not patent documents but mainly scientific publications. In a subsequent investigation [7] on bio-engineering, Narin and colleagues found that the 'time delay' for references to scientific literature in patents is comparable to the same in scientific publications.

Van Vianen, Moed & van Raan [15] studied chemical engineering in an international context, as well as Dutch technology as a whole, in particular chemical engineering and electrical engineering and electronics. They, too, based their work on US patents and uncovered not only valuable results on important patent citation characteristics, but also a series of methodological problems. For instance, in the US patents used, a strong language barrier works to the disadvantage of all non-English literature.

Coward and Franklin [2] also tried to establish a particular relationship between science literature and patent literature in order to determine 'cross-overs' between science and technology. The suggestion was made that the most productive processes for identifying potential profit-yielding areas of scientific research for industrial technology, involve the identification of areas in which researchers were able to produce frequently scientific publications *and* patents.

Recently, Rabeharisoa [10] studied the role of scientific articles published by inventors. She focused on the rather narrow field of French fuel cell R&D in order to assess the contents of papers as well as patents. Her case analysis illustrates the intermediate dynamic function in the science and innovation complex by pointing to the crucial role of, for instance, technical papers. Rabeharisoa's contribution clearly demonstrates that it will be an unsuccessful task to differentiate scientific from technological researchers or experts, at least in fuel cell R&D.

The aim of this study is to investigate one specific aspect of the S&T interface: *inventor-author relations*. We chose *laser applications in medicine* as our subject area. Grupp & Schmoch [6] present a discussion about this choice, together with an overview of related (sub)fields in laser R&D. Laser applications in medicine cover therapeutic as well as diagnostic instruments. It is clear that this field has strong relations with physics (in particular, applied optics) as well as with medical fields. A very important and early application of lasers in medicine concerns eye surgery. With laser-ophthalmoscopy, retinal detachments (in particular in the case of diabetes patients) are treated ('spot-welding' of the retina). One of the first publications is the work of Smart [11]. A year earlier Goldman [3] had published another biomedical application of lasers, the treatment of skin cancer. Shortly after these first medical applications, Thorp [12] foresees in his review book considerable advances in the field. And indeed, in 1985 the number of patents in laser medicine was almost an order of magnitude larger than ten years before (Grupp & Schmoch [6]). For publications we find an even stronger increase (Van Vianen & van Raan [14]). Therefore, we conclude that laser medicine R&D is a dynamic and relatively young area.

4.2 Method and Techniques

4.2.1 Main lines

The core of our empirical work is a set of 30 patent applications (European Patent Office, EPO) and in addition all publications of the inventors for the period 1980-1989, as far as covered by the Science Citation Index (SCI) of the Institute of Scientific Information (ISI). The total number of publications is 1053.

Our study includes four different approaches. First, we divide our set of 30 randomly selected patents (European Patent Office, priority years mainly 1986-1988) into three subsets according to the number of references in the patent search reports to non-patent (mostly scientific) literature ('examiner-given' NPL references; see, e.g. Grupp & Schmoch [5,6], and van Vianen, Moed, van Raan [15]). In order to investigate whether the number of NPL references in patents represents a measure of 'science intensity', we analyze for each patent general publication characteristics of the inventor-authored publications, in relation to the three NPL-based patent subsets.

The second approach is a 'refinement' of the first: we now focus, for each patent, on only one specific inventor-authored publication (if present) that can be regarded as most closely related to the subject matter of the patent. The choice was made by expert opinion. Again, for this specific type of inventor-authored publications we searched for bibliometric characteristics in relation to each of the three NPL-based patent subsets. The first and the second approach thus emphasize the background of an invention, in particular its science intensity.

Third, we study the 30 patent documents in full length. With the help of experts in the field, each patent was characterized according to its dependence on recent scientific results. Further, characteristics of the patent claims, in particular the complexity of the commercial issues to be protected, the complexity of the invention, the size of the team of inventors, as well as the multiplicity of possible applications have been investigated in relation to this dependence on recent scientific results. Again, these characterizations were analyzed in relation to the three patent subsets.

Fourth, we drop the distinction of the three NPL-based patent subsets (and, therefore, leave the relation with science intensity). We now focus again on the total oeuvre of the inventor-authors, in order to find characteristics of patent-related publications, in particular their time-dependent behavior in relation to patent priority years. Thus, this fourth approach emphasizes the scientific side of an invention. We analyzed two particular characteristics of anticipated importance: the share of university-company co-operative publications in the total amount of inventor-authored publications, and the degree of inventor co-authorship, i.e., the number of inventors per patent involved in research publications.

4.2.2 Data collection

The creation of our publication database is based on three sets of ten patents each, applied for at the EPO. The composition of a set is determined by the number of non-patent literature (NPL) references in the patents. One set contains patents with no NPL references, the second contains patents with exactly one NPL reference, and the third contains patents with more than three NPL references. Of these 30 patents, the 72 inventors' names were searched for in the 1980 to 1989 CD-ROM versions of the Science Citation Index (SCI) of the Institute for Scientific Information (ISI). Our search resulted in 1053 publications with 2006 addresses. We found that the inventors of only four of the 30 patents were not represented in the SCI. In all these four cases it concerned a patent with only one inventor. For another four patents we found only one publication of a (co-) inventor (group) in the SCI. In all other cases there were three or more publications (not necessarily authored by all inventors of a patent). Two co-inventor groups even produced more than 200 publications in the ten-year period studied.

At the level of the individual inventors, we found that 58 out of 78 inventors published at least once in a journal covered by SCI (74.4%). We emphasize that we focus on publications covered by the SCI. Therefore, in this study we do not include publications that might be covered by other databases than SCI, such as EMBASE or MEDLINE. For a more extensive and detailed study of the inventors' scientific oeuvre, the use of these databases may be of importance (Van Vianen and Van Raan [14]).

4.2.3 Details of the four different approaches

As discussed above, our study includes four approaches of analyzing the data. The first is a comparison of general characteristics of publications authored by the inventors involved in the three subsets of patents.

Our hypothesis is that the subset containing patents with the most NPL references is the most science intensive, and that this science intensity should be reflected by specific characteristics of the inventor-authored publications for each patent subset. Possible characteristics are ranking lists of the publishing authors, the affiliations of the authors, the journals selected by them, and the publication types used. The list of journals for each subset was examined in more detail by taking the nature of the journals into account. For this purpose, the 'journal level' classification by Noma [9] has been used (level 1 refers to technology-oriented applied journals, whereas level 4 represents the other side of the 'spectrum', the most basic journals). Moreover, a preliminary citation analysis has been performed for the inventors' publications (1980-1989), in order to analyze their characteristic (scientific) impact for each of the three different patent subsets and thus to validate this characteristic against the original division of the patents on the basis of the number of NPL references.

In a second approach, we analyze those publications that are probably most closely related (MCR) to the subject matter of the patents. We selected, if present, one publication per patent. We assume that these MCR publications represent the scientific counterparts of the patented inventions. We identified per patent in each set these MCR publications by comparing names of co-inventors with names of co-authors, as well as addresses of patent's applicants with affiliations of the authors, the year of patent application with publication year, and by comparing the subject matter of the patent with the title of publications. We determined for each of the three patent subsets the average number of references given in the publications and investigated whether the group of MCR publications with the highest average number of references corresponds to the subset of patents with the highest number of NPL references. Furthermore, the MCR publications were subject to a citation analysis.

In the third approach, experts in the field of laser medicine analyzed the 30 patents carefully in terms of their dependence on (basic) scientific results by studying the full patent documents. In this expert analysis important aspects of the patented invention were taken into account, such as complexity of the invention (i.e. how many claims), the size of the inventor team and the multiplicity of possible applications. With the help of these data, scores for closeness to and dependency from recent scientific results were given for each patent. We emphasize that the experts did not know about

NPL references and our grouping of the 30 patents on the basis of these NPL references.

In our fourth approach, all inventor-authored publications of the three patent subsets were taken together, and grouped again, but now according to several criteria referring to hypotheses on characteristics of inventor-authored publications.

We studied the distribution of publications closest to the patents over a ten year period (1980-1989) around the date of patent application, in order to unravel publication strategies of inventors.

We formulated two hypotheses about inventor-authored publications. To test these hypotheses, we dropped the subset division as based on the number of NPL references in the patents and adopted a new division based on bibliometric characteristics of the publications themselves. Herewith, we shifted our point of reference from patents to publications. By comparison of the distribution around the date of patent application of those publications assumed to be 'closest to the patents' with the distribution of the whole set of inventors' publications, our hypotheses were tested.

We made the following assumptions: (1) the higher the number of *co-inventors* also being co-authors of a publication, the more this publication will be *related* to the patent; (2) *cooperative* publications of universities and companies (or hospitals) indicate a higher degree of application-oriented research, indicated by patent relatedness.

In order to test the first assumption, we divided the complete set of publications into three subsets. The number of co-inventors (CI) being co-authors (CA) of a publication is indicated by the 'CICA score'. Publications of which more than 50% of the co-inventors of a patent appeared to be co-authors, are classified with CICA=1; publications with exactly 50% of the co-inventors are (co-) authors with CICA=2; and publications with less than 50% of the co-inventors as (co-) authors with CICA=3. The time trends (percentage per year) of each CICA subset was studied in a ten-year period around the date of patent application.

For the second assumption, we focus on those publications, having both a university and a company as a corporate address, indicating a co-operation of these two types of institutions (U&C). We suppose that these U&C publications are also closely related to the patents, assuming that companies in most cases are applicants of the patents concerned. Again we made a trend analysis (percentage per year) with this set of U&C publications.

4.3 Results and Discussion

4.3.1 First approach: general bibliometric characteristics

Our first approach concerns characteristics of inventor-authored publications, in particular the relation of specific characteristics with the three patent subsets, which are based on the number of NPL references in the patents. Our assumption was that the characteristics to be analyzed may indicate a relation with 'science intensity' of the patented invention, i.e., its relatedness with academic research. Our goal is to find empirical evidence in favor of or against the claim that the number of NPL references in the patents is a measure of science intensity.

More or less to our surprise, the great majority of the inventors appears to publish at least once in journals covered by the SCI. Basic data of the inventor-authored publications are given in Table 4–1. We show for each patent of the three subsets the number of inventor-authored publications, the number of inventors being authors, and the number of citations to inventor-authored publications in the (sample-) years 1983, 1985, 1987, and 1989.

Table 4–1 Number of publications (Npu), number of inventors (Nin), number of inventor being author (Nau), and number of citations (Nci)

Patent set 1 (NPL=0)					Patent set 2 (NPL=1)					Patent set 3 (NPL=3)				
Ptnr	*Npu*	*Nin*	*Nau*	*Nci*	*Patnr*	*Npu*	*Nin*	*Nau*	*Nci*	*Patnr*	*Npu*	*Nin*	*Nau*	*Nci*
1.01	6	1	1	7	2.01	0	1	0	0	3.01	0	1	0	0
1.02	8	3	2	5	2.02	1	3	3	0	3.02	3	4	1	8
1.03	23	1	1	33	2.03	20	2	2	19	3.03	17	3	3	5
1.04	16	3	2	31	2.04	15	3	2	6	3.04	12	2	2	13
1.05	17	4	4	63	2.05	1	3	1	0	3.05	13	2	2	24
1.06	37	4	3	18	2.06	56	2	2	9	3.06	216	5	4	580
1.07	5	2	1	18	2.07	269	3	3	252	3.07	16	7	3	29
1.08	0	1	0	0	2.08	139	5	5	149	3.08	16	1	1	2
1.09	3	2	1	0	2.09	1	1	1	0	3.09	1	1	1	0
1.10	59	4	4	97	2.10	0	1	0	0	3.10	83	3	3	186
	174	25	19	272		502	24	19	435		377	29	20	847

The numbers of inventor-authored publications seem to differentiate somewhat between the three sets. A closer look at the data, however, points out that the differences are primarily due to one or two teams per set. Set 2, for instance, owes its highest number of publications almost exclusively to the productivity of one group of co-inventors, namely the inventors of patent nr. 2.07. On the other hand, two groups of co-inventors in this patent set do not have any publication in SCI-covered journals (patents 2.01 and 2.10), whereas in patent set 1 only one group is found without publications covered by SCI (patent 1.08). Moreover, in patent set 1 all other nine patents have inventors who produced at least three publications, whereas in patent set 2, there are three patents with inventors who produced only one (SCI) publication. This may illustrate that these basic figures are not sufficient to bring out the possible differences between the three patent sets. Thus, mere numbers of inventor-authored publications *do not discriminate significantly* between degrees of patent science intensity, as far as this latter characteristic is measured by the number of NPL references in the patents.

Also the number of inventors being authors does not yield notable differences between the three sets. In all three sets, about 75%of the inventors (of lasers in medical applications) are authors of at least one publication covered by the SCI.

In order to find further evidence in favor of or against the presupposition (Narin et al.[8]) that the number of NPL references in patents indicates a degree of science intensity, we performed a preliminary publication citation analysis. Citations given in 1983, 1985, 1987, and 1989 (sample years) to *publications* of the inventors (1980 - 1989) of each patent were counted in the SCI (last column, Table 4–1). The overall figures show remarkable differences between the three patent sets. At first sight, patent set 3 appears to be the most 'science intensive', in terms of received citations by inventor-authored publications. Only half of that number is received by patent set 2, and just a fourth by set 1. But again, the high number of citations for patent sets 2 and 3 are primarily due to 'outliers', two inventor teams (with this we mean the inventors of one specific patent) of patents 3.06 and 3.10. In set 2, as much as five inventor teams receive no citations at all in the years included in our analysis. Those inventor teams with the most publications, also have the most citations (patents 2.07 and 2.08). If we neglect the outliers in sets 2 and 3, it becomes problematic to decide which set represent the most science intensive one, as far as it concerns received citations. It is certainly not set 2, but even set 3 shows, if we only neglect patentnr. 3.06, a less favorable 'spread' in number of citations than set 1.

Apparently, more information is required about inventor-authored publications in order to decide whether a specific patent set is more science intensive than another. Therefore, we made an analysis of further characteristics of each set. For this purpose a bibliometric profile was created for the publications of each set. These profiles have been developed as a standard bibliometric tool of the Leiden group. They contain the following information elements: (1) author names of publications; (2) corporate addresses of authors; (3) type of articles; and (4) journals used for publications.

In this study we focus on (2) and (4). In Table 4–2 we present the results for the addresses (the first 10 in ranking). Comparing the ten most frequently occurring addresses of the three sets, we could hardly find any characteristic differences. Each set has about five university addresses in its top 10. Assuming that publications with a university address should be considered as more basic research oriented, we notice that the publications in all three sets appear to have the same character. However, it should be noted that the SCI coverage tends to be less complete at the applied side of the journal spectrum, so that there might be a bias against publications with company addresses.

Table 4–2 Addresses of inventor-authored publications (Npu= number of publications)

Patent Set 1

Rank	*Npu*	*Institute*
1	29	INST OPHTHALMOL, LONDON, GREAT BRITAIN
2	22	UNIV LONDON, LONDON, GREAT BRITAIN
3	9	MESSERSCHMITT BOLKOW BLOHM GMB, MUNICH, FED REP GER
4	9	STADT KRANKENHAUS MUNCHEN, MUNICH, FED REP GER
5	9	UNIV MUNICH, MUNICH, FED REP GER
6	8	NIPPON HOSO KYOKAI, TOKYO, JAPAN
7	7	MED UNIV LUBECK, LUBECK, FED REP GER
8	6	NATL CANC CTR, TOKYO, JAPAN
9	6	UNIV CHICAGO, CHICAGO, USA
10	5	COLUMBIA UNIV, NEW YORK, USA

Patent Set 2

Rank	*Npu*	*Institute*
1	141	HARVARD UNIV, BOSTON, USA
2	89	UNIV HEIDELBERG, HEIDELBERG, FED REP GER
3	31	CORNELL UNIV, NEW YORK, USA
4	20	RIVERSIDE RES INST, NEW YORK, USA
5	16	UNIV FREIBURG, FREIBURG, FED REP GER
6	14	THOMSON CSF, PARIS, FRANCE
7	13	MAX PLANCK INST STROMUNGSFORSC, GOTTINGEN, FED REP GER
8	12	UNIV ARIZONA, TUCSON, USA
9	8	UNIV CALIF LAWRENCE LIVERMORE, LIVERMORE, USA
10	5	EUROPEAN MOLEC BIOL LAB, HEIDELBERG, FED REP GER

Patent Set 3

Rank	*Npu*	*Institute*
1	135	UNIV PENN, PHILADELPHIA, USA
2	45	TNO, RIJSWIJK, NETHERLANDS
3	18	ERASMUS UNIV, ROTTERDAM, NETHERLANDS
4	18	UNIV CALIF LAWRENCE LIVERMORE, LIVERMORE, USA
5	16	MAX PLANCK INST BIOPHYS CHEM, GOTTINGEN, FED REP GER
6	14	ALCON LABS INC, FT WORTH, USA
7	13	ONCOGEN, SEATTLE, USA
8	13	UNIV GRENOBLE 1, GRENOBLE, FRANCE
9	12	HOKKAIDO UNIV, SAPPORO, JAPAN
10	12	NEW YORK UNIV, NEW YORK, USA

Table 4–3 Journals used by inventors for publications(Npu= number of publications)

Patent Set 1

Rank	*Npu*	*Journal*
1	21	LASERS IN SURGERY AND MEDICINE
2	10	PROCEEDINGS OF THE SOCIETY OF PHOTO-OPTICAL INSTRUMENTATION ENGINEERS
3	7	EUROPEAN UROLOGY
4	6	APPLIED OPTICS
5	6	BRITISH JOURNAL OF OPHTHALMOLOGY
6	6	GASTROINTESTINAL ENDOSCOPY
7	6	JOURNAL OF THE OPTICAL SOCIETY OF AMERICA
8	5	EXPERIMENTAL EYE RESEARCH
9	5	TRANSACTIONS OF THE OPHTHALMOLOGICAL SOCIETIES OF THE UNITED KINGDOM
10	4	INVESTIGATIVE OPHTHALMOLOGY & VISUAL SCIENCE

Patent Set 2

Rank	*Npu*	*Journal*
1	79	JOURNAL OF INVESTIGATIVE DERMATOLOGY
2	50	CLINICAL RESEARCH
3	22	LASERS IN SURGERY AND MEDICINE
4	15	BERICHTE DER BUNSEN GESELLSCHAFT FUR PHYSIKALISCHE CHEMIE
5	14	HUMAN GENETICS
6	14	JOURNAL OF THE AMERICAN ACADEMY OF DERMATOLOGY
7	13	ARCHIVES OF DERMATOLOGY
8	12	AMERICAN JOURNAL OF OPHTHALMOLOGY
9	12	PHOTOCHEMISTRY AND PHOTOBIOLOGY
10	11	APPLIED PHYSICS B-PHOTOPHYSICS AND LASER CHEMISTRY

Patent Set 3

Rank	*Npu*	*Journal*
1	29	CHEMICAL PHYSICS LETTERS
2	25	JOURNAL OF CHEMICAL PHYSICS
3	21	BIOPHYSICAL JOURNAL
4	13	BIOCHEMISTRY
5	13	EXPERIMENTAL HEMATOLOGY
6	11	APPLIED OPTICS
7	11	APPLIED PHYSICS B-PHOTOPHYSICS AND LASER CHEMISTRY
8	11	CYTOMETRY
9	10	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA
10	9	ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY

So far, we must conclude that the differences in all the above analyzed characteristics between the three patent sets appear to be not significant enough to support the claim that the number of NPL references in a patent represents a measure of science intensity in one way or another.

Our bibliometric profiles, however, give us another information element: the journals (as far as covered by the SCI) in which the publications appeared. Similar to the addresses, we focused our attention to the top-10 in the ranking list for each patent set (Table 4–3).

At first sight, these rankings tend to be rather inaccessible. It is hard to judge which journal list belonging to one of the three patent sets, is more or less basic-oriented than another. Therefore, we used the 'level indicator' of journals (Noma [9]). With help of experts' judgments, a level (1 to 4) is assigned to each journal of the SCI. To mention the extreme cases, level 1 is attached to journals considered as 'very applied' or 'technology-oriented', whereas journals of level 4 are considered to have a typical basic research character. With help of this simple typology, we were able to characterize the 'basic versus applied nature' of the journals, and therefore in first approximation also of the inventor-authored publications concerned. We counted per patent set the number (frequency) of publications with a specific level. These data are represented by a frequency distribution in Figure 4-1.

The results of this analysis are quite remarkable. As illustrated by Figure 4-1, a clear shift of the frequency distribution from set 1 to 3 is visible. Whereas the inventor-authors of patent set 1 have chosen for more technologically oriented journals, the inventor-authors of set 3, use much more basic journals. The inventor-authors of set 2 are just in between. With the knowledge that patent set 1 is based on patents with no NPL references, set 2 on patents with just one, and set 3 on patents with more than three NPL references, our journal-level finding appears to be the *only* bibliometric indicator, as far as investigated in this study, *in favor* of the claim that the number of NPL references is an indicator of 'science intensity' and that therefore the patents of set 3 are the most science-intensive (i.e., the most basic research-oriented).

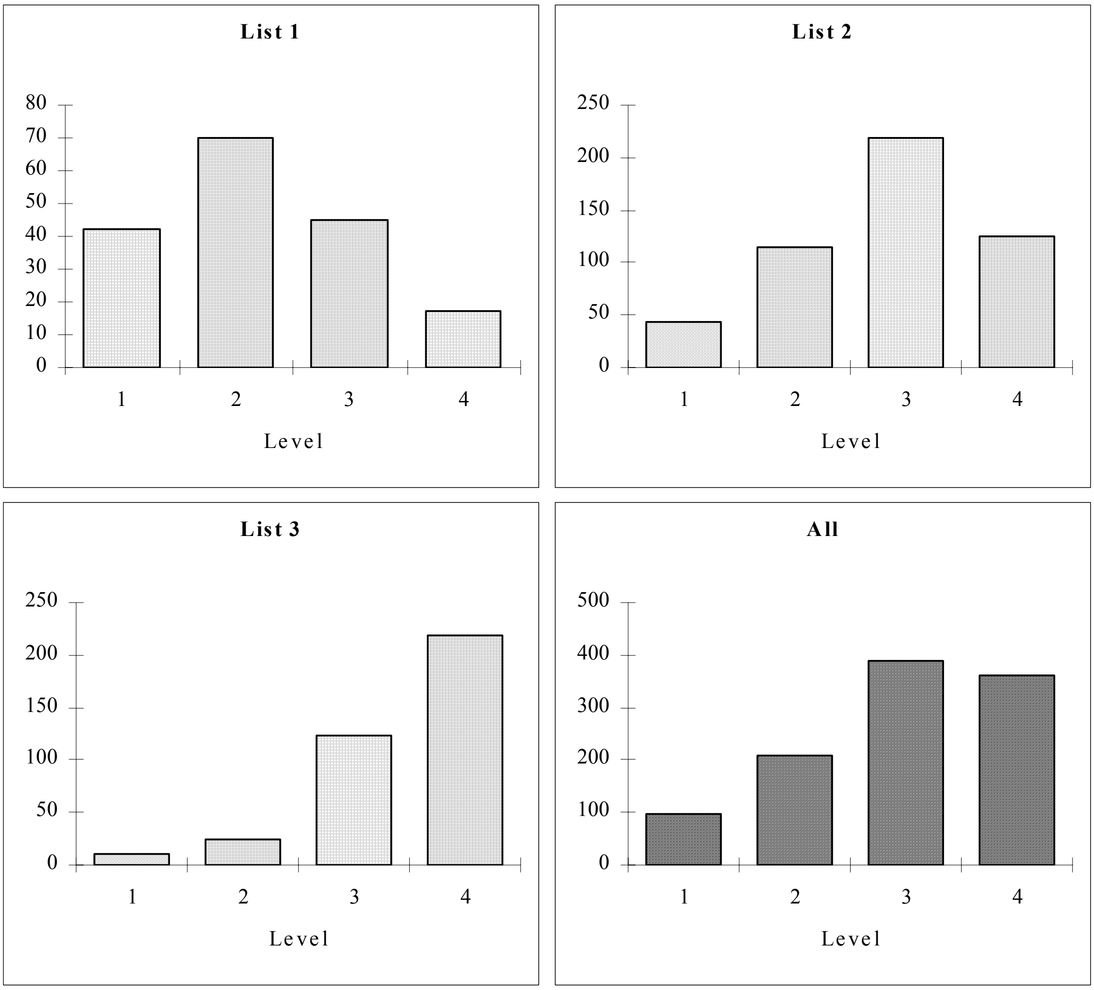

Figure 4-1 Distribution of inventor-authored publications over 4 journal levels

4.3.2 Second approach: scientific counterparts of patents

In the above discussed first approach the 'science side' of a patented invention was represented by the complete publication oeuvre of inventors (1980-1989, and as far as represented by the SCI). This approach has the disadvantage of taking all scientific activities of inventors into account, also those that are probably not (closely) related to the patented work. As we are interested in the science intensity of the particular inventions, this may be a drawback. The work leading to a patented invention may only be just part of the inventors' research activity at that time. Therefore, we tried to identify for each patent one inventor-authored publication that may be regarded as being most closely related to the patent work. In other words, we try to find a *scientific counterpart* of each patent. For 18 of the 30 patents, we identified (using expert opinions) such a most closely related (MCR) publication. We arranged these MCR publications according to the three original patent sets, and analyzed their specific characteristics.

In order to structure our analysis, we first formulate two hypotheses. The first one is that the number of references in MCR publications and the number of NPL references in the patents are highly correlated. This would imply that the average number of references of MCR publications related to patents in set 1 is significantly lower than the average number of references of MCR publications related to patent set 3. Apparently, as we can see in Table 4–4, this is indeed the case.

Table 4–4 Number of references (Nre) in MCR publications per patent (Nmcr) for each patent set

Patent set 1		*Patent set 2*		*Patent set 3*	
1.01	14	2.01	n a	3.01	n a
1.02	6	2.02	5	3.02	n a
1.03	3	2.03	n a	3.03	14
1.04	0	2.04	n a	3.04	4
1.05	7	2.05	n a	3.05	8
1.06	0	2.06	12	3.06	47
1.07	n a	2.07	18	3.07	36
1.08	n a	2.08	11	3.08	9
1.09	n a	2.09	n a	3.09	n a
1.10	20	2.10	n a	3.10	14
Nmcr	7		4		7
Nre	50		46		132
Mean	7.1		11.5		18.9
STD	6.9		4.6		15.0

In both patent sets (1 and 3), we found a MCR publication for seven patents. The average number of references in the case of set 3 is significantly higher. Moreover, in set 1, two MCR publications appeared to have no references. With respect to set 2, we found for only 4 patents a MCR publication, which may indicate a less science-intensive character than, for instance, set 1. On the other hand, these four publications have a significantly higher number of references per paper than the seven publications of set 1 (11.5 against 7.1). With respect to our first hypothesis, we must conclude that there is only a weak support for the claim that the number of NPL references indicates science intensity of patents. Our findings that set 2 only has four patents with a MCR publication against seven in both set 1 and set 3, is also not in favor of our first hypothesis.

Our second hypothesis is the assumption that science intensity of patents may be related to the scientific impact of the corresponding MCR publications. Taking into account that the application of most of the patents was relatively recent, we assessed the short-term impact, using a 'three-year' citation count window, see Table 4–5.

Table 4–5 Number of citations (Nci, three year citation-count-period) to MCR publications per patent for each patent set

Patent set 1		*Patent set 2*		*Patent set 3*	
1.01	11	2.01	n a	3.01	n a
1.02	7	2.02	2	3.02	n a
1.03	3	2.03	n a	3.03	17
1.04	0	2.04	n a	3.04	2
1.05	7	2.05	n a	3.05	8
1.06	1	2.06	5	3.06	18
1.07	n a	2.07	12	3.07	2
1.08	n a	2.08	4	3.08	3
1.09	n a	2.09	n a	3.09	n a
1.10	15	2.10	n a	3.10	3
Nmcr	7		4		7
Nre	44		23		53
Mean	6.3		5.8		7.6
STD	5.0		3.8		6.6

As compared with the results of Table 4–4, the differences between the three patent sets are even much less significant. Still, the average number of citations per MCR publication is slightly higher in set 3 than in sets 1 and 2 (7.6 against 6.3 and 5.8). This implies that the (short-term) impact of (applied) research related to patents is more or less the same for all three sets. Thus we find little support for the assumption that the science intensity of patents (as measured by the number of NPL references) is correlated to the (short-term) impact of MCR publications.

So far, the first two approaches yield the following results concerning the science intensity of patents:

(1) the *number of NPL references* in patents correlates significantly with the basic versus applied *nature of journals* used by inventors for their publications, i.e., with the type of inventor-authored publications in terms of basic versus technology-oriented research;

(2) the *number of NPL references* in patents correlates only weakly with the *number of references* in those inventor-authored publications, that are *most closely related* with the patented invention (MCR publication).

As mentioned before, it is a drawback of the first approach that the whole oeuvre (1980-1989) of inventors is included. Hence, publication characteristics may apply to an inventor primarily as a scientific author, rather than as an inventor. In the second approach, however, the selection of only one publication that can be considered as being most closely related to a patent, reduces the numbers so radically that problems with statistical significance arise. This raises the question how meaningful the distinction of the three patent sets according to the NPL references really is.

4.3.3 Third approach: expert opinions

The findings of the third approach, based on expert opinions on each patent, are as follows (the experts involved did not know about of the number of NPL references).

Table 4–6 Principle properties of patent documents and expert assessment

Patent set 1						Patent set 2					Patent set 3				
Nr	*EA*	*Ncl*	*Npg*	*Ncc*	*Nin*	*EA*	*Ncl*	*Npg*	*Ncc*	*Nin*	*EA*	*Ncl*	*Npg*	*Ncc*	*Nin*
1	1	27	7	5	1	1	14	12	2	1	3	47	11	2	1
2	3	67	26	1	3	1	15	4	5	3	5	78	17	3	4
3	1	11	6	2	1	1	9	7	2	2	3	20	6	1	3
4	2	8	9	3	3	1	6	6	2	3	1	46	14	1	2
5	1	35	11	2	4	2	7	4	2	3	1	49	14	1	2
6	1	10	4	3	4	3	21	8	3	2	5	40	24	2	5
7	1	14	4	2	2	4	11	7	1	3	1	7	9	1	7
8	3	11	4	1	1	4	3	3	4	5	4	48	27	2	1
9	1	7	3	3	2	1	6	6	2	1	1	19	10	2	1
10	3	10	15	3	4	1	10	5	1	1	3	10	6	2	3
Mean	1.7	20.0	8.9	2.5	2.5	1.9	10.2	6.2	2.4	2.4	2.7	36.4	13.8	1.7	2.9
STD	0.9	17.9	6.6	1.1	1.2	1.2	5.0	2.4	1.2	1.2	1.6	20.9	6.7	0.6	1.9

EA= expert assessment of science involvement (scale 1 to 5); Ncl= number of claims; Npg= number of pages; Ncc= number of classification codes; Nin= number of inventors; STD= standard deviation)

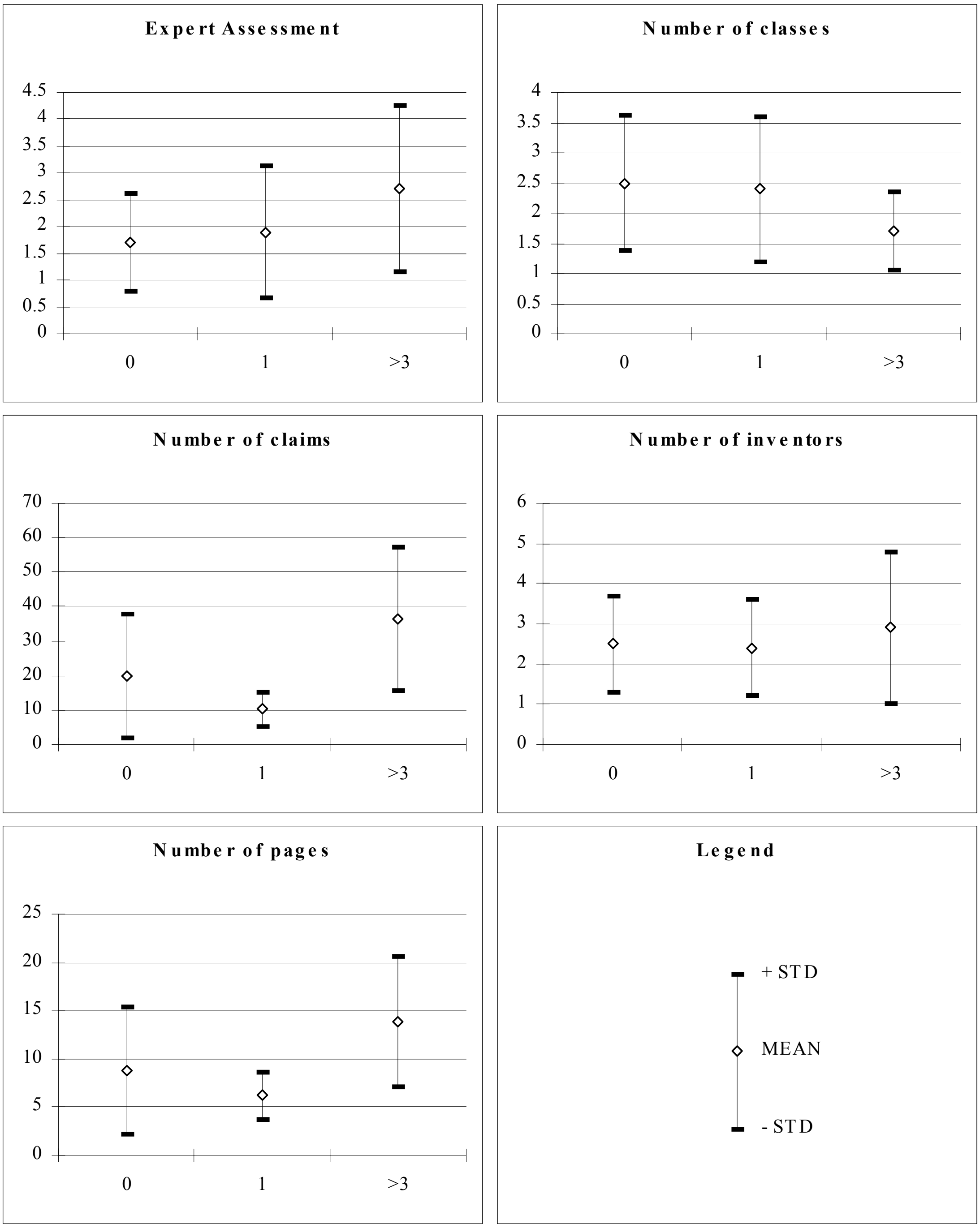

Figure 4-2 NPL-references vs. 5 other indicators

As shown in Table 4–6 and Figure 4-2, the expert assessment of the 'science involvement' of the patents correlates with the number of NPL references in the expected direction. The more NPL references patents contain, the higher the science intensity scores are. However, there is a large variance between the sets of ten patents so that the differences in experts' scores, again, are *not* significant. The large standard deviation may be due to the low number of patents in the sets, but the experts think that also a larger sample probably will not yield significant differences. The experts indicated to us that, generally, frequent resort to scientific papers is a necessity in certain patent documents for other reasons than science intensity. This remark stimulated us to investigate some other features of these patents. What are these possible other reasons? Patent examiners have to check the novelty of an invention claim by claim. If there are many claims in a patent application, then the legal situation requires that the examiners give more references to earlier patents and/or to scientific publications (i.e. proofs). Indeed, as Figure 4-2 shows, patents with more than three NPL references are characterized (with a large variance) by more claims than other patents. This means that there is another reason than only science intensity which requires more frequent NPL references. A third issue is the *complexity* of the invention which, according to the experts, can be measured by the length of the document (the description of the invention and its background). Patents with more than three NPL references appear to be longer and, according to the experts, more complex. Difficult deliberations with regard to topics such as tissue properties, radiation exposure, and 'half-width' of energy in skin, take place in laser medicine. Sometimes cumulated citations to physiology, radiation biology and cancer research are given together with those to laser or atomic physics just because the invention covers more than laser treatment of one organ or one specific application.

There are two other reasons for a patent examiner to refer in the search report of an individual patent application to scientific literature: complex contents and a complex array of legal claims. Therefore, we conclude that the whole field of laser medicine is science dependent.

The above conclusion is further supported by the observation that those patents with more than three NPL references are somewhat more central in terms of classification (i.e., multiple classification less likely; see Figure 4-2). They might be more complex in contents, more complex in legal terms and more science-intensive, but often they are classified in hierarchically higher patent classes because their specificity is not as clear as in other cases. The three sets of patents do not differ in size of inventor teams, i.e., those with more NPL references, are *not* related to larger teams.

Table 4–7 Common features of 30 patents in Lasers in Medicine

Expert assessment of science involvement	*N Ncl*	*N Npg*	*N Ncc*	*N Nin*	*NPL refs.*
Closest (*N*=5)	36.0 (27.0)	15.6 (9.3)	2.4 (1.0)	3.6 (1.5)	2.6 (1.4)
Close (*N*=9)	22.3 (19.7)	9.8 (6.5)	2.0 (0.8)	2.7 (1.1)	1.4 (1.7)
Not close (*N*=16)	17.8 (13.6)	7.6 (3.5)	2.3 (1.2)	2.3 (1.6)	1.3 (1.4)

For Ncl, Npg, etc., see legend Table 4–6. Closest: expert scores 4 and 5 in Table 4–6; Close: expert scores 2 and 3; Not close: expert score 1 (standard deviations between parentheses).

Table 4–7 presents an 'inverted view'. Here we re-group the 30 patents according to the expert scores for science intensity: one set with patents scoring 1, another set scoring 2 or 3, and a third set scoring 4 or 5. Again, it is demonstrated that NPL references as well as the legal and cognitive complexities (number of claims and pages) increase with science intensity as defined by the experts. However, none of these relations is very significant, the most suggestive one being the length of the document. In this representation team size does seem to correlate with science intensity, in contrast with our findings in Table 4–6.

The above observations confirm that science intensity is an intrinsic feature of a technology field *as a whole* (Grupp and Schmoch [5], p.90) and not as much an *individual* property of a patent document within that field. Therefore, if a sample of patents contains *many* NPL references, then an individual patent in that sample with little or even without such references may nevertheless also be influenced by science, just as patents with many NPL references may have a rather remote science link (Table 4–6). Yet, patents differ by type of journal used by the inventors for their scientific publications: 'more complex' inventions with a larger number of NPL references are linked to inventor-authored publications in more basic journals, as opposed to 'less complex' inventions with no or a few NPL references. These 'less complex inventions' are related to applied or technology-oriented journals. Insofar as a journal indicates a type of research, the three original patent subsets are linked to either basic or more applied/ technology-oriented research. Still, all subsets are linked to science in general, as indicated by the extent of scientific activity of inventors and the expert assessments. We conclude that the number of NPL references seems to correlate more with *type* of research.

4.3.4 Fourth approach: time trends in inventor-author relations

4.3.4.1 Two basic indicators

In our fourth approach we return to the data set used for the first approach: the complete scientific oeuvre of inventors of 1980 - 1989 (as far as published in SCI-covered journals). We will make an attempt to develop further indicators for tracking down the inventors' publication activity period before a patent application. In this way, we eventually may be able to make a better selection from an inventor's whole oeuvre of those publications relevant to the patent.

For this analysis, our earlier identification of just one MCR publication per patent is dropped. As mentioned above, all inventors' publications are used. We hypothesize that two specific bibliometric characteristics of publications would indicate 'closeness' to patents. Based on these characteristics, the database is divided into one subset which can be regarded as a collection of publications being relatively close to the patented invention and one subset being less close, to be used as a test group. Both subsets are subject to a trend analysis over a ten-year period (1980-1989).

One additional remark must be made about the database. The priority years of the patents involved in our analysis are not the same. Most of the patents, however, have 1986 or 1987 as priority year. In order to avoid inaccuracies introduced by mixing different time trends, we restricted the data of our analysis only on the basis of patents with priority year 1985, 1986 or 1987. This reduces the number of patents to 22, and the number of publications (by inventors of the thus selected patents) to 581, with 1117 addresses.

The two specific bibliometric characteristics (in quantified form: indicators) are the following.

(1) Number of co-inventors being co-authors

This first characteristic indicating 'closeness' to a patented invention is the relative number of co-inventors being also co-authors of a publication. We divided all inventor-authored publications into three subsets, on the basis of the CICA score as discussed in Section 2.3: (1) publications with more than 50% of co-inventors as co-authors (CICA=1); (2) publications with 50% of co-inventors as co-authors (CICA=2); (3) publications with less than 50% of co-inventors as co-authors (CICA=3).

The distribution of publications in each CICA set over a period of ten years is shown in Table 4–8, smoothed per 2 years (i.e., numbers of 2 successive years are added and divided by 2) to reduce annual fluctuations. In addition, the relative activity per two-year block (i.e., the percentage of publications per block relative to the total production in the whole period) is given. Figure 4-3 shows these percentage distribution results.

Table 4–8 Publication trend per CICA-group (1980-1989)

	CICA=1		CICA=2		CICA=3		All	
Period	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
80-81	6.5	8.2	4.0	4.5	46.5	11.2	57.0	9.8
81-82	7.0	8.9	3.5	4.0	57.5	13.9	68.0	11.7
82-83	6.0	7.6	8.5	9.7	55.0	13.3	69.5	12.0
83-84	4.5	5.7	11.5	13.1	43.5	10.5	59.5	10.2
84-85	10.0	12.7	14.0	15.9	36.5	8.8	60.5	10.4
85-86	**12.5**	**15.8**	**15.0**	**17.0**	**46.0**	**11.1**	**73.5**	**12.7**
86-87	**13.0**	**16.5**	**11.5**	**13.1**	**44.0**	**10.6**	**68.5**	**11.8**
87-88	8.5	10.8	8.5	9.7	30.0	7.2	47.0	8.1
88-89	4.0	5.1	6.0	6.8	25.0	6.0	35.0	6.0

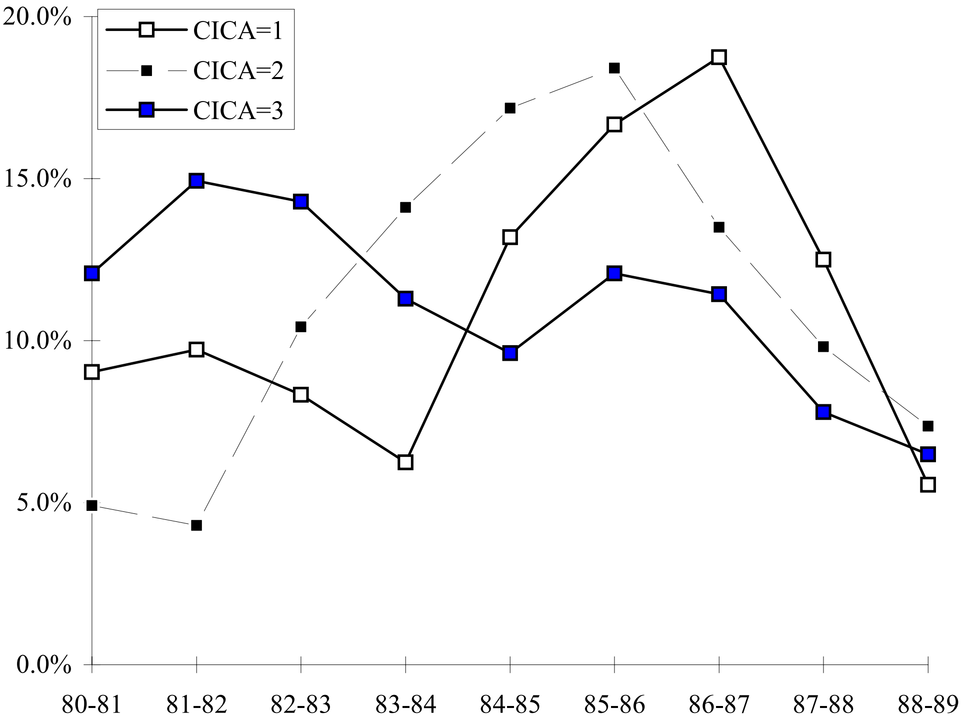

Figure 4-3 Distribution of inventor-authored publications over a 10 years period

In the sets of 'CICA=1' and 'CICA=2' publications, a significant increase of publication activity starts around 1984/1985 and 1983/1984, respectively. Afterwards (around 1986/1987), it returns to the level of before 1984. The number of 'CICA=3' publications, however, is rather stable throughout almost the whole period (around

10%), but decreases in the most recent years. The set of all publications shows the same trend.

As the activity increase of 'CICA=1' and 'CICA=2' culminates just before the year of patent application period, we may conclude that the inventor-authored publications in these subsets, indeed, are most closely related to the patent work. This would mean that around the date of patent application, co-inventors also increase their co-activities in scientific research, as far as reflected by publication numbers (in SCI-covered journals). With the above approach, we can not decide whether these inventor-written papers do reflect research activities parallel to the patent work and therefore might be characterized as the 'scientific counterparts of the patent', or cover more application-oriented research which in fact *precedes* the patent. The most plausible explanation, however, is that there is R&D activity on a specific topic in laser medicine, the work becomes successful and more and more results are published in (applied) scientific articles. Meanwhile, also technological application reaches the stage of concrete possibilities and materializes in a patent application. Thus, we choose for the 'parallel' process in which scientific and technological work go hand in hand. Nevertheless, as far as published knowledge concerned, we conclude that publication data precedes patent data.

(2) Co-operation of universities and companies

The second characteristic indicating 'closeness of publications to patents', is the co-operation of universities with companies. From the patent data, we learned that in most cases the applicant of patent is a company. On the other hand, we also found that most of the inventors are affiliated to universities. Therefore, it is interesting to investigate trends of university-company collaboration. For this analysis, we composed two new subsets of the inventor-authored publications. One subset contains publications with both a university and a company address, and the other contains all other publications. As in the previous analysis, both subsets are subject to a trend analysis over a ten- year period. The results are represented in Table 4–9 and Figure 4-4.

Table 4–9 Publication trend for university-company (U&C) co-operation (1980-1989)

	U&C		Other		All	
period	*N*	*%*	*N*	*%*	*N*	*%*
80-81	6.0	8.5	51.0	10.0	57.0	9.8
81-82	2.5	3.5	65.5	12.8	68.0	11.7
82-83	2.5	3.5	67.0	13.1	69.5	12.0
83-84	5.0	7.0	54.5	10.7	59.5	10.2
84-85	7.0	9.9	53.5	10.5	60.5	10.4
85-86	**10.0**	**14.1**	**63.5**	**12.5**	**73.5**	**12.7**
86-87	**13.0**	**18.3**	**55.5**	**10.9**	**68.5**	**11.8**
87-88	10.0	14.1	37.0	7.3	47.0	8.1
88-89	7.0	9.9	28.0	5.5	35.0	6.0

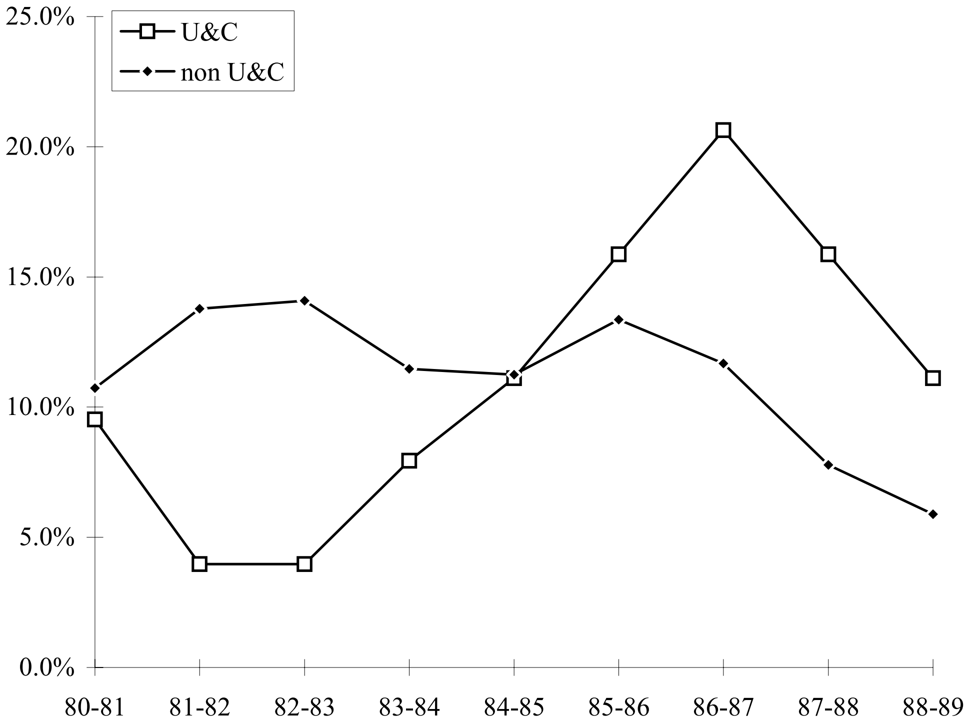

Figure 4-4 Distribution of inventor-authored publications over a 10 years period, university/company co-operations vs. others

Although the total number of publications involved is slightly lower than in the CICA analysis, the distribution is remarkably similar. From about 1984, there is a sharp activity increase of university-company co-operation (again, as far as reflected by publications covered by the SCI), culminating around the years of patent application, 1986/1987. Afterwards, the U&C activity decreases to an average number of papers per year. A notable dip is found in 1980/81, long before the increase, and we are still uncertain about the cause. The trend of the other publications (i.e., with only a university, or only a company as an address, or any other address) fluctuates, like the overall trend, around the average, again with a small decrease in the two most recent years.

We conclude that we found a second bibliometric indicator to define 'patent-related publications'. We showed that this second bibliometric indicator also illustrates an increase of R&D activities in the period around patent application, in the same way as the CICA indicator.

4.3.4.2 Patent application vs. publishing

In addition to these results, we focus on a delicate issue concerning publishing inventors. It is clear that a new invention must be original in order to be patented. It is therefore not surprising that a patent application can only be accepted if nothing has been published about the invention up to the day of application. Not by others, nor by the inventors themselves. An actual publication about the work described in a patent, and, thus, with the same innovation disclosures, can therefore only be published after the application date of the patent involved.

Taking this consideration into account, one should conclude that the publications responsible for the increase of co-activity of inventors ('CICA=1' publications) and the increase of co-operation between universities and companies (U&C publications) around the date of patent application, do not include the papers which do actually disclose the invention. In that case, a patent application would never be granted for lack of novelty.

It illustrates, however, once more that patents applications and scientific papers should be considered as two different components of the same (in broader terms) R&D output. In our sample, these two components seem to be complementary.

4.4 Conclusions

The number of non-patent literature (NPL) references in laser medicine patents correlates significantly with the degree of 'appliedness' of the inventor-authored publications in the period around the patent priority year. Furthermore, we found a correlation, although not strong, between the number of NPL references in the patents and the number of references in those inventor-authored publications that are most related to the patent work.

What do these two findings mean? We think that they do *not* prove that the number of NPL references in patents are a measure, for an *individual* patent, of the 'science intensity' as such. But, rather, a technological field or specialty is science-intensive as such, or, in other words, science intensity is an *intrinsic property* of certain (sub-) fields of technology (Grupp & Schmoch [5, pp.90-98]). In our opinion, the first finding shows that there are patents in a specific field of technology (in this case: laser medicine) that are typically *technological* in nature (e.g., new instrumental developments). For the R&D work involved, both the 'science side' (publications) as well as the 'technology' side (patents) thus have a more applied character, resulting in the use of typical applied journals, and less (or none) NPL references in patents. The conclusion could be that *if* patent work is more related to applied research, the necessity of the patent-examiners to explicitly list NPL references in the patent decreases. Thus, inventions with technology orientation can well be science-intensive.

The second finding is quite interesting in the light of the above explanation. The scientific work related to the more typically technological patents (having less or none NPL references) also seems to have less references in the publications. This, however, is in fact the other side of the same coin: Narin [7] found that the more a publication is applied in nature, the less references are given (in the SCI data).

In conclusion, we found that less or no NPL references in patents is not necessarily an indicator of a lesser science intensity of the individual patents (since, for example, the number of inventor-authored publications and the received citations *do not* discriminate between patents with less or more NPL references!), but an indicator of the more technological nature of individual patents. Thus, patent documents with no or only one NPL reference cannot be regarded as significantly less 'science-intensive' as such (but probably: less *basic-research*-intensive) than those with many NPL references.

The findings based on expert opinions on the individual patent documents, indicate that the entire R&D field of laser medicine depends on scientific progress. Those patents containing more NPL references are often more complex, i.e., include more claims, but are not necessarily more science intensive. They may, however, be more basic-research intensive. This meshes with the expert opinion that a complex set of legal claims corresponds to a more general description of possible commercial applications.

If we drop the original distinction of three NPL-based patent subsets (and, with that, the relation with NPL-based indicators of science intensity) and focus on the total scientific oeuvre of all patent inventors, we find further important characteristics.

First, there is a significant increase of inventor co-authorship in the period before the patent priority year. Second, the data show a significant increase of university-company collaboration (as far as reflected by our bibliometric method) again in the period before the patent priority year. It is plausible, however, that both phenomena are related: if part of the inventors' team is affiliated to a company, then an increasing university-company collaboration implies a similar trend in inventor co-authorship.

Bibliography

[1] Carpenter, M.P., M. Cooper, and F. Narin, 'Linkage between basic research literature and patents', *Research Management*, 13 (1980), 30-35

[2] Coward, H.R., and J.J. Franklin, 'Identifying the science-technology interface: Matching patent data to a bibliometric model', *Science, Technology and Human Values*, 14 (1989), 50-77

[3] Goldman, L., Medical and Surgical Users for the Laser, *New Scientist*, 21 (1964), 284

[4] Grupp, H. (ed.), *Dynamics of science-based innovation*, Heidelberg: Springer Verlag, 1992

[5] Grupp, H. and U. Schmoch, 'Perception of scientification of innovation as measured by referencing between patents and papers. In: H. Grupp (ed.), *Dynamics of Science-based Innovation* (Springer Verlag, Heidelberg, 1992), pp. 73-128

[6] Grupp, H. and U. Schmoch, 'At the crossroads in laser medicine and polyimide chemistry - Patent assessment of the expansion of knowledge'. In: H. Grupp (ed.), *Dynamics of Science-based Innovation* (Springer Verlag, Heidelberg, 1992), pp. 269-301

[7] Narin, F. and E. Noma, 'Is technology becoming science?', *Scientometrics*, 7 (1985), 369-381

[8] Narin, F., M. Rosen, and D. Olivastro, 'Patent citation analysis: New validation studies and linkage statistics'. In: A.F.J. van Raan, A.J. Nederhof, H.F. Moed (eds.), *Science and technology indicators*. Leiden: DSWO Press, 1989, 35-47

[9] Noma, E., Subject Classification and Influence Weights for 3000 Journals (Computer Horizons Inc., Cherry Hill, NJ, 1986)

[10] Rabeharisoa, V., 'A special mediation between science and technology: when inventors publish scientific articles in fuel cells research'. In: H. Grupp (ed.), *Dynamics of Science-based Innovation* (Springer Verlag, Heidelberg, 1992), pp. 45-72

[11] Smart, D., *New Scientist*, 26 (1965), 570

[12] Thorp, J.S., *Masers and Lasers: Physics and Design* (Macmillan, London, 1967)

[13] Van Raan, A.F.J. (ed.), *Handbook for Quantitative Studies on Science and Technology*. Amsterdam Elsevier Science Publishers, 1988

[14] Van Vianen, B. & A.F.J. van Raan, 'Knowledge expansion in applied science: A bibliometric study of laser medicine and polyimide chemistry'. In: H. Grupp, *Dynamics of Science-based Innovation* (Springer Verlag, Heidelberg, 1992), pp. 227-267

[15] Van Vianen, B., H.F. Moed, and A.F.J. van Raan, 'An exploration of the Science Base of Recent Technology', *Research Policy*, 19 (1990), 61-81