Contents

Introduction

This document contains a draft version of the final report for the study concerning "Future priorities for Community research based on bibliometric analysis of publication activity," in view of the Five-year Assessment (1999-2003) of Community research activities. This study is carried out by EPEC (European Policy Evaluation Consortium) for DG Research under Lot 1 of the Framework Contract for ex-ante evaluation (reference BUDG-02-01-L1), and more precisely by Technopolis France in cooperation with the Observatoire des Sciences et des Techniques – the French national S&T indicator institute.

It is set out in 4 further sections

- Background of objective of the study
- Methodological approach
- Presentation of results
- Conclusion and recommendations

The report has two annexes:

- Literature review
- Detailed overview of data cleaning procedure

1 Background and objectives to the study

1.1 Background to the study

The Five-year Assessment[1] of Community research activities covers ex post and ex ante evaluative aspects. It enables the achievements and impact of past and ongoing activities to be evaluated and appraised, contributes to making implementation of these activities more transparent and plays a significant role in defining the future direction of research, technological development and demonstration activities. The ex ante role of The Five-year Assessment is underlined by the exercise being a compulsory stage in the preparation of subsequent Framework programmes.

The Five-year Assessment is carried out by independent highly qualified experts, supported by separate analyses carried out by independent external specialists with experience in socio-economic, scientific and technological issues, state-of-the-art survey techniques and research programmes.

The assessment is carried out in a broader policy context set by the Heads of State and Government at the Lisbon Summit in March 2000 which underline the need to prepare the transition towards a competitive, dynamic, knowledge-based economy and to develop a European Research Area.

1 Article 6 (OJ L 232 of 29.8.2002 pp. 3 and 36) and Article 8 (OJ L 294 of 29.10.2002 pp 4 and 46).

Within such a perspective, a quantitative evaluation of Framework Programmes is important. Bibliometric indicators form in principle a good means to check the value of a research programme for the tax payer: to say it simply if a researcher who receives public money publishes, that means s/he has produced an output; if moreover a publication is well-cited, that means that, generally, the scientific results are read and found relevant to be taken into account in other publications. Understanding the output of the Framework Programme – and this study is a first in its kind – and its potential scientific impact, should constitute an important element in monitoring the programme in the future.

1.2 Bibliometrics and priority setting[2]

Ever since bibliometric indicators exist, their potential for use in the policy process has been investigated… and questioned. Whereas numbers of publications are a generally accepted indicator to represent the productivity of (basic) researchers, in countries such as the Netherlands, citation analysis even had very early on been proposed as a means of evaluating their performance and, eventually, as a basis to distribute research funding. The "impact factor" is generally accepted to evaluate the relevance of a publication, and on higher levels can be used in order to analyse the relative position of research entities or organisations. Patent analysis is used by industry as a tool to support technology watch, and by countries to rate their technological performance; and in the second half of the nineties the bibliometric analysis of norms and standards was added to more traditional publications and patents analysis. With computer power and adapted software increasingly available, and increasing acceptance of network-based methods,[3] bibliometrics seems to regain attention, as testified by the recent WREN workshop, co-organised in Brussels by the European Commission and US DoE.

On a European level the use of bibliometrics in the policy making process has been increasingly on the agenda in recent years, as testified for instance by the exercise of "Mapping excellence in research and technological development (RTD)" which aimed "at providing useful intelligence allowing greater advantage to be taken of the potential of excellent competencies available across the EU" – a validation study however showed that mapping excellence by using bibliometrics alone was too ambitious an objective.[4]

Also for the present study it should be borne in mind that, from a theoretical perspective, bibliometric analysis does not serve priority setting directly, and that this task is even more difficult an issue in the context of the Framework Programme – for at least two reasons.

First, as the terms "bibliometrics" indicates, publications should be available to be analysed. As a matter of fact, the potential use of bibliometrics differs enormously between areas of research, both in terms of scientific or technological disciplines and

2 See also the Annexe A for a quick literature review on bibliometrics and impact assessment.

3 Grounded in the Sociology of innovation, the use of network based approaches for scientometrics can be traced back to the early 1980s, cf Callon, Michel, John Law and Arie Rip (eds.), 1986, *Mapping the Dynamics of Science and Technology*, Houndsmill: MacMillan Press.

4 V. Charlet, B. de Laat & A. Lantrain, 2003, Mapping of Excellence in Immunology RTD, Brussels: European Commission.

in terms of types of research (basic, applied…). Differences are generally so great that a harmonised approach by definition is impossible. This seems especially true in the case of the Framework Programme since it covers research of all sorts and from various disciplines. For instance, under the IST priority, certain types of research (e.g. software) may well "vanish" in a standard bibliometric analysis, whereas less interdisciplinary parts of the FP may be overrepresented. Applying bibliometric techniques to the European Framework Programme is therefore particularly challenging.

Second, it should be observed that, within priority setting, bibliometrics, which is basically an evaluative, *ex post*, methodology, is not used in a "linear" way – i.e. it does not provide decision makers directly with options for choice, but may support them in making eventual choices.[5] Bibliometrics is one tool and source of information amongst often many others.

1.3 Objectives of the study

The aim of the study is to **devise and implement a methodology to exploit and if possible enhance existing data sources** on publications resulting from Community funded research and for its analysis. This was done with a view on the identification of future priorities for Community funded research, based on analysis of scientific and other publication activity. The study supports the Five-year Assessment of Community research activities (1999-2003).

The analysis examines patterns of publication activity and the impact of these publications, using bibliometric techniques.

The study was designed with the view of helping policy makers to answer the following question: what are the priorities for Community funded research on the basis of

- the perceived quality and impact of Community funded research compared with that of non-Community-funded research?
- trends in terms of publishing activity for Community funded research?
- comparison with the body of knowledge on research publishing in general, especially in terms of the overall European picture and internationally?

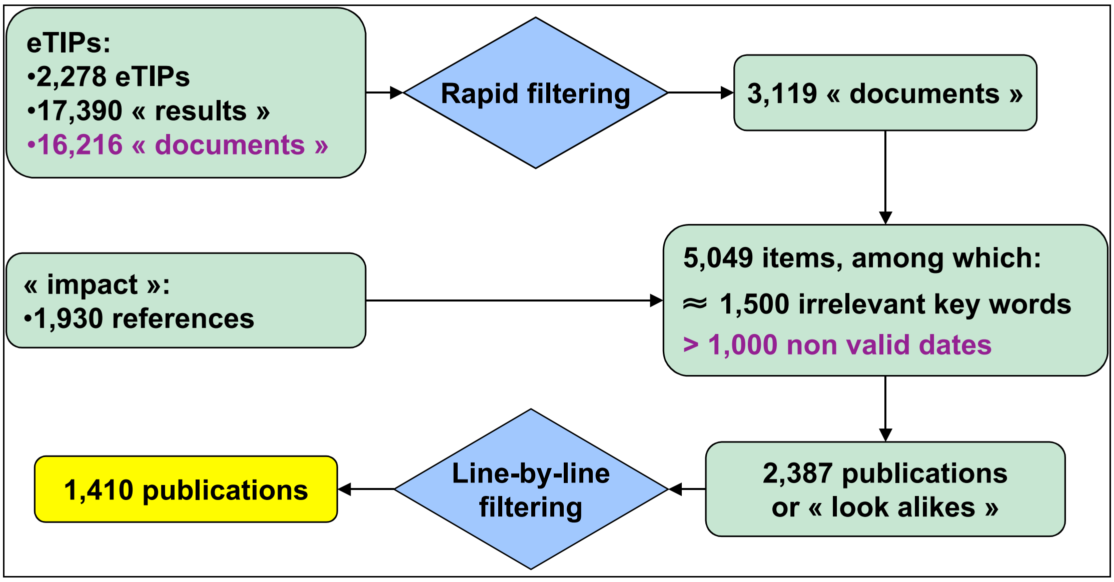
As we suggested quite early on, "priorities" refer not so much to substance (which areas should be precisely promoted…) at this stage: this seems too premature, given the fact that the present study is a first one in its kind, starting from information bases of poor quality. Much more, the study aims to describe the methodological and monitoring priorities for the use bibliometric techniques to assess the impact of the Framework Programme, on the basis of real empirical findings and results.

5 Attempts to provide intelligence for priority setting have however been made since the mid-1980s, cf. Courtial, Jean-Pierre, 1990. Introduction à la Scientométrie. De la bibliométrie à la veille technologique, Paris: Anthropos Economica; Courtial, J.-P., et al., 1993. The use of patent titles for identifying the topics of invention and forecasting trends, Scientometrics, vol.26, n°2, p.231-242.

2 Methodological approach: from raw to ready-to-use data

This chapter describes the successive steps that were followed to obtain data which could be processed to obtain indicators, starting from raw data provided by the Commission. Exhibit 1 provides a comprehensive overview of the cleaning process.

Exhibit 1 Overview of the data cleaning process



2.1 Description of databases received from the Commission

Technopolis France received two databases from the European Commission, from which to extract bibliometric indicators:

- a file labelled "publications_extraction3.zip" on 31/08/2004 (hereafter named "eTIPs database")
- a file labelled "EVAL base de référence 12-11.zip" on 27/08/2004 (hereafter named "impact database")

2.1.1 The "eTIPs database"

The eTIPs database is an MS Access database, derived from an extraction of data in more than 2,000 projects' eTIPs.[6] Corresponding projects are at different levels of finalisation and not all of the eTIPS have been approved. As detailed below, the database includes information on project outputs of various kinds.

It covers 2,144 projects and 2,278 eTIPs. The difference in numbers of projects and eTIPs is most probably due to revised eTIPs not overwriting earlier versions.[7]

[6] "electronic Technology Implementation Plans" are to be filled in by all participants in FP projects, to provide information on projects' outputs.

[7] See B. de Laat & F. Martin (2003), *Methodological analysis of TIP as a tool for impact assessment and evaluation*. Brussels: European Commission DG RTD.

eTIPs contain, in principle, all individual project results. As each project – and therefore each eTIP – may have led to more than one result, the database contains 17,390 results altogether.

Additionally, each of these results in turn point at different "outputs" divided into four different families:

- Corresponding thematic field(s): 45,397 "subject descriptors" are mentioned altogether to qualify the results
- Corresponding market applications: 28,236 applications are contained in the database
- Corresponding patents: 6,567 patents are included
- Corresponding documents: 16,216 items are mentioned

Academic publications are part of what are named "documents"; since the latter can be of many types (conference presentations, software products, web pages…), a significant part of the (manual/visual) data processing has consisted of spotting publications amongst these 16,216 items.

This appeared to be quite difficult due to the heterogeneity of the responses. For instance, the field "document type" which would have normally enabled to sort all items into publications, conferences, software documentation… was filled inconsistently. Not only were there lots of possible spellings for one given type ("Publication", "publication", "peer-reviewed paper", "paper", "journal paper"…), there were also many cases in which the entry consisted of the reference itself ("publication" or "report title"…).

After a first, rapid, but incomplete, filtering (i.e. suppression of all items easily spotted as irrelevant: patents, reports, PhD theses…) 3,119 items were kept for further treatments.

2.1.2 The "impact database"

The second dataset was an MS Excel spreadsheet containing 2,206 responses to a recent FP3/FP4 impact questionnaire.

Among other questions, respondents were asked to indicate the first and the second best publications derived from their project. Out of the 4,412 possible answers, 1,930 references were provided.

By construction, all responses were supposed to correspond to academic publications – many of them however appeared not.

2.2 Cleaning of the resulting database

2.2.1 Semi manual cleaning

The fusion of both databases resulted in a set of 5,049 documents, among which were still remaining other documents than academic publications.

The following steps were run semi manually:

- deleting all items (around 1,500) containing irrelevant key words ("symposium", "presentation", "congress"…)
- deleting all items (more than 1,000) which did not contain any valid data (and could therefore not be matched with the OST database).

This resulted into a list containing 2,514 items (publications or "look-alikes")[8]. Once double items were erased, the list contained 2,373 items.[9]

2.2.2 Manual line-by-line filtering

Another line-by-line filtering phase was necessary before calculating indicators.

The OST staff therefore erased all remaining irrelevant items, that is:
- in many cases, non publications items or incomplete references (journal title was often missing)
- more rarely, references to articles published in other journals than those present in the ISI database at the reference year (2000).[10]

Submitted articles were kept together with already published articles.

This finally lead to a total of **1,410 publications** (858 from the impact database and 552 from the eTIPs database).

2.3 Relevance of results

Considering the above detailed data characteristics, to date the results of this study **cannot be viewed as representative** for the total of publications produced by the research funded through the Framework Programme. More precisely:
- The sample of publications given above can only be expected to be *a small part* of all papers published thanks to FP funding
- It is impossible to estimate the size of the whole and thus the proportion taken by this part on the whole
- Different parts of the programme have different publication patterns (see below)
- It is not possible to assess whether there was a bias in the cleaning stages towards certain types/subjects of research or certain periods of time

Therefore, the only remaining possibility to document this question of representativeness is to assess the proportion of FP funded **projects** covered by this study. Some information is presented further down in this report (see section 4); as explained above, this only gives an approximate indication about the covered proportion of **publications**.

---

8 The resulting Excel® file was communicated electronically to the European Commission (N. Reeve, A. Silvani) on 14 September 2004 for inspection.

9 An error-checking process was then run by the Commission services on the remainder (5 049 – 2 373 = 2 676 items). It revealed that 69 publications were unintentionally removed from the sample during this semi-automated filtering phase or could have been identified on a pure (but unaffordable) line-by-line approach.

10 See Annexe B for a detailed description of this process.

The results of the present study should not be taken at face value to assess the scientific impact of the Framework Programme. The real result of this study is that **it proves the feasibility of the proposed methodological approach**.

3 Bibliometric results

3.1 General description of the publication set

The publication set obtained can be characterised as follows:

- 709 different journals are represented
- 8 journals out of these 709 get more than 10 references each. The most recurrent of them – J. Biol. Chem. – gets 32 "hits" (see Exhibit 2).
- The maximum yearly number of articles was reached in 2002. The past seven years (1998-2004) correspond to more than 50 articles per year (see Exhibit 3).
- This period represents the vast majority of the publications: previous papers (from 1991 to 1997) only represent 6% of the total number (see Exhibit 4).

Exhibit 2 Titles of the 30 most recurrent journals in the final set

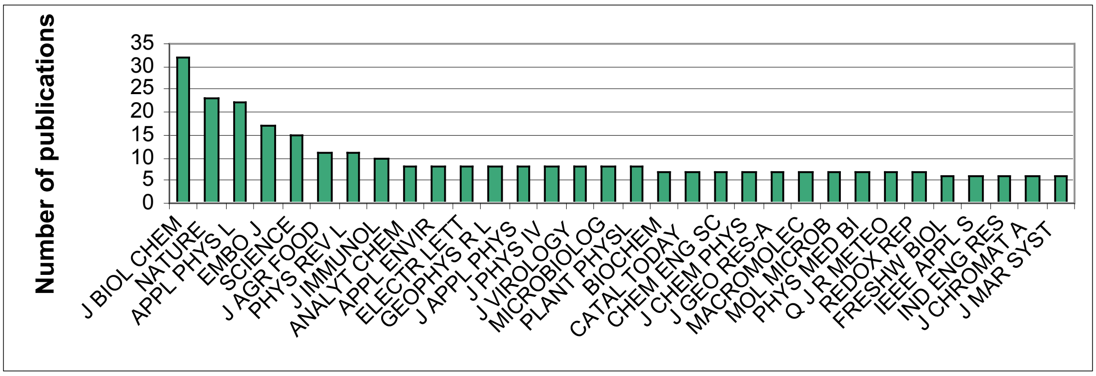


Exhibit 3 Number of articles present in the final set, by publication year

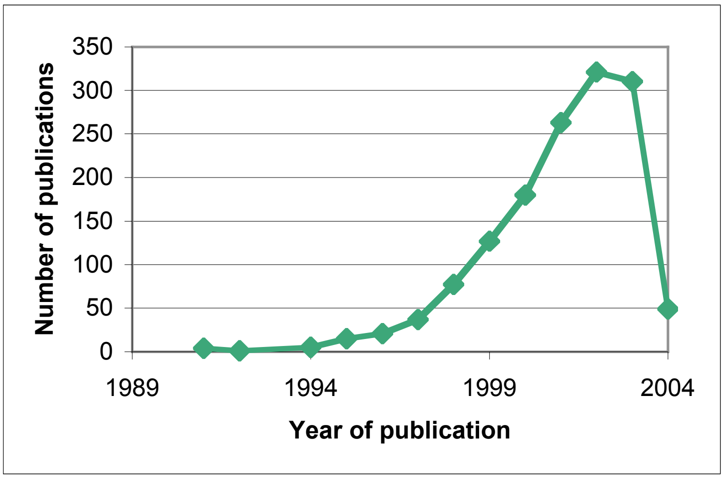
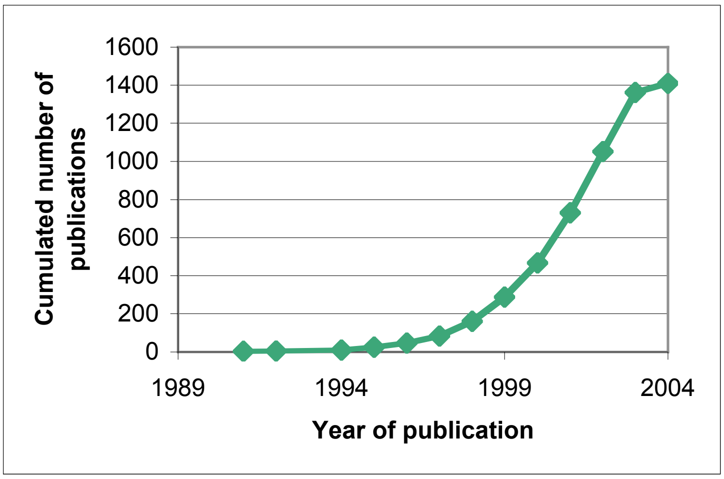


Exhibit 4 Cumulated number of published papers



3.2 Focus on the most represented research programmes

3.2.1 Overall description

16 RTD programmes appear as the funding sources of the research projects corresponding to these 1,410 papers, out of which 6 get more than 100 references each (see Exhibit 5).

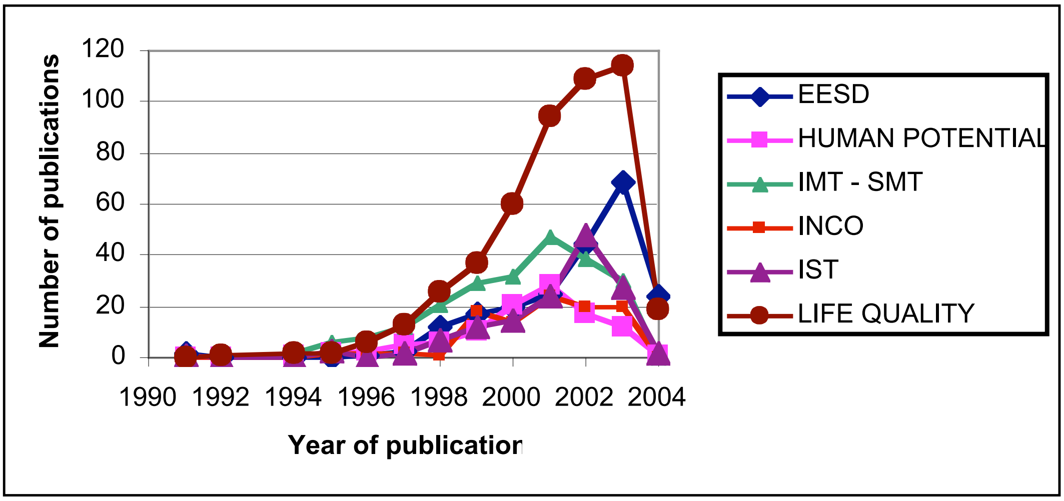
Exhibit 5 Number of analysed articles, by funding programme

Programme	Framework programme	Number of articles
QoL (Quality of Life)	5th FWP	483
IMT – SMT	4th FWP	227
EESD (Sustainable Development)	5th FWP	218
IST (Information Society)	5th FWP	141
Human potential	5th FWP	106
INCO (International Cooperation)	4th FWP	103
Growth	5th FWP	50
FP5-EAECTP C (Euratom)	5th FWP	19
INCO 2 (International Cooperation)	5th FWP	13
Fission	6th FWP	13
BRITE/EURAM 3	4th FWP	12
FAIR	4th FWP	10
ESPRIT 4	4th FWP	5
MAST 3	4th FWP	4
ENV 2C	4th FWP	3
Autre	na	2
NNE-JOULE C	4th FWP	1
Total		**1 410**

Note: there was a bias in the original data since projects from “older” specific programmes (such as Biomed 2, Biotech 2, Brite/Euram 3, Esprit…) were sometimes attributed to the most recent programme that could be associated to it (QoL, IMT-SMT, IST…). The later programmes are therefore overrepresented as compared to the older ones.

Although these funding programmes range from the 4th to the 6th FWP and therefore cover a long time period, publication distributions nearly have the same profile (see Exhibit 6).

Exhibit 6 Yearly number of published articles for the 6 most represented programmes (1991-2004)



3.2.2 Programme-by-programme analysis

As could be expected, publications corresponding to different specific programmes – and presumably to different research fields – are not expected to be published in the same journals. The following exhibits give detailed distributions of journals for the 5 most represented programmes.

Exhibit 7 The 30 most recurrent journals related to the programme QoL

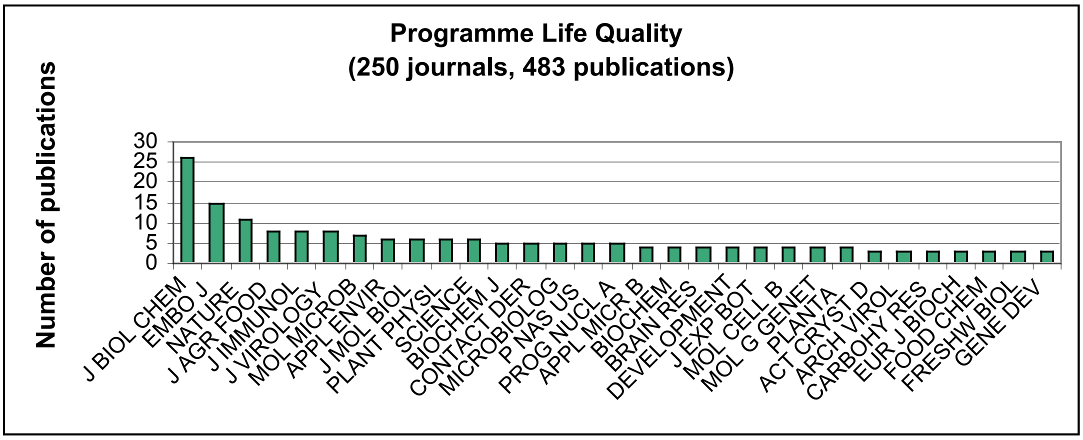


Exhibit 8 The 30 most recurrent journals related to the programme IMT - SMT

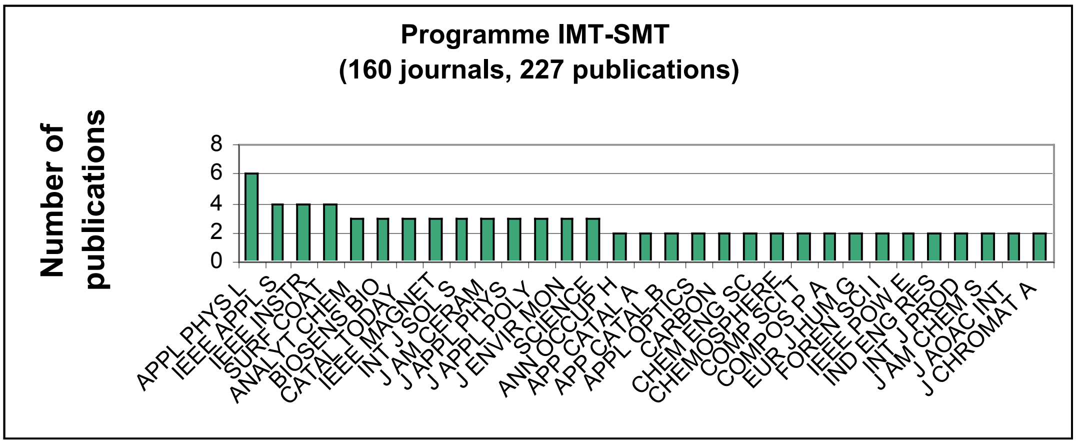


Exhibit 9 The 30 most recurrent journals related to the programme EESD

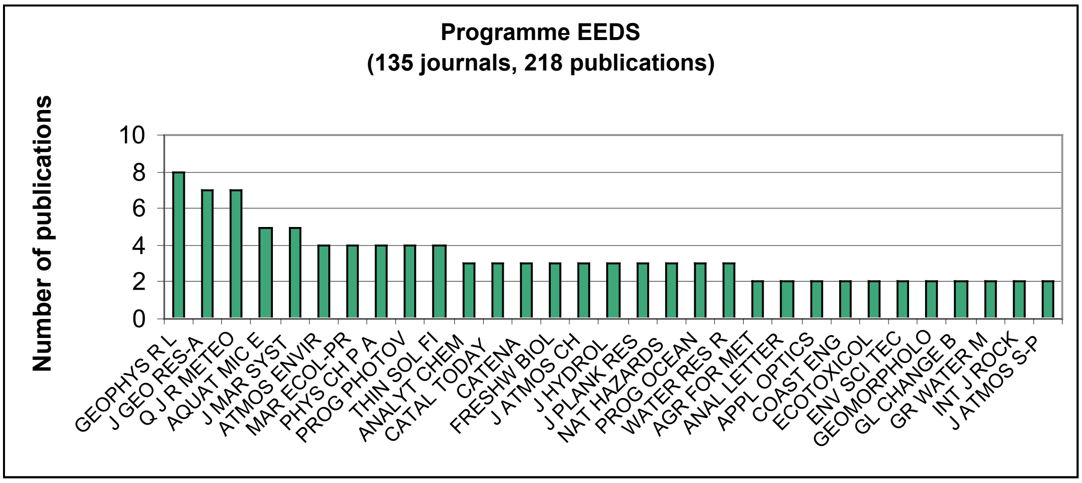


Exhibit 10 The 30 most recurrent journals related to the programme IST

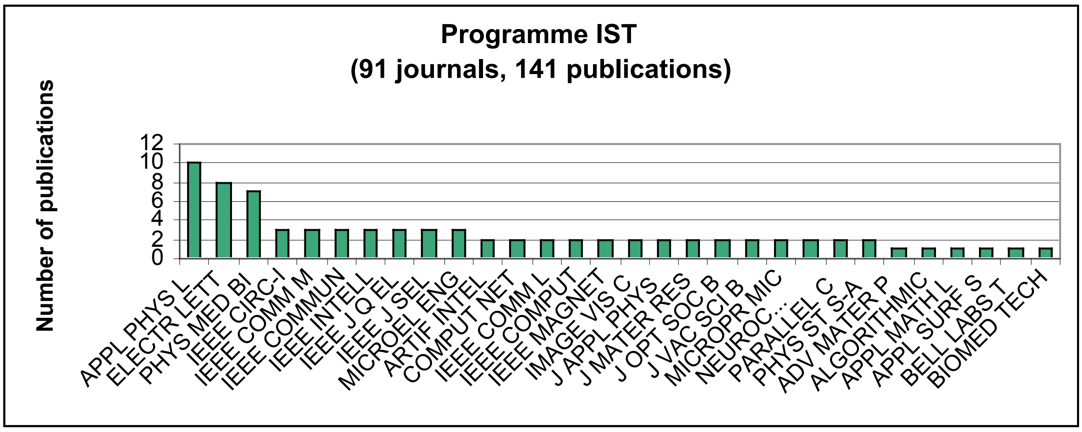
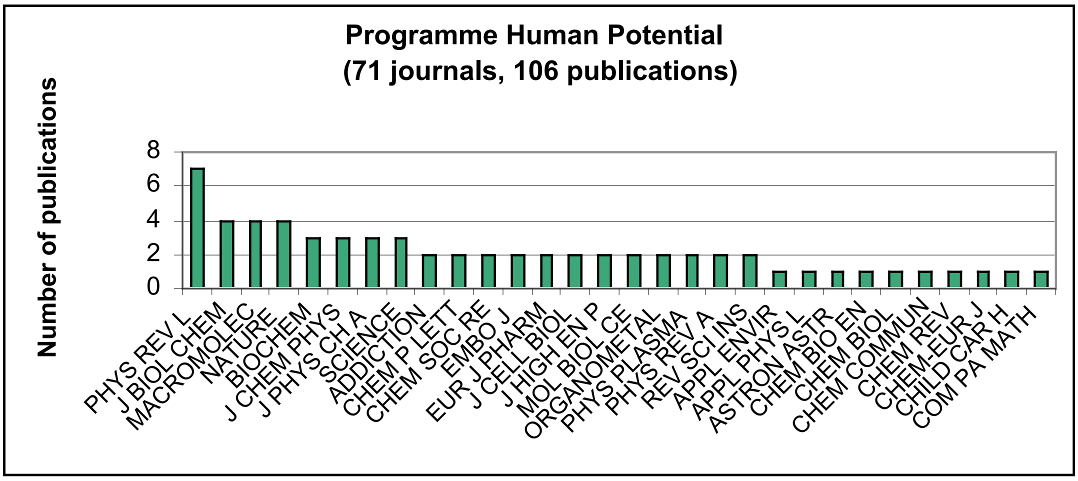


Exhibit 11 The 30 most recurrent journals related to the programme Human Potential



3.3 Estimating the impact of journals

3.3.1 Definition of the bibliometric notion of impact

The rationale of this study is to try and measure the impact of RTD programmes, using as a proxy the impact of academic journals in which corresponding articles were published.

In bibliometric terms, the notion of impact responds to a precise definition, namely **the number of citations received by a given article**. This definition relies on the implicit assumption that citations are a sound representation of the extent to which a given publication has contributed to the progress of knowledge.

Although it is theoretically possible to calculate the impact of each article, the available data do not allow for such a detailed analysis: since publications' authors, titles and references could not be isolated, these publications were assumed to correspond to the "average article" of the journal in which they were published.

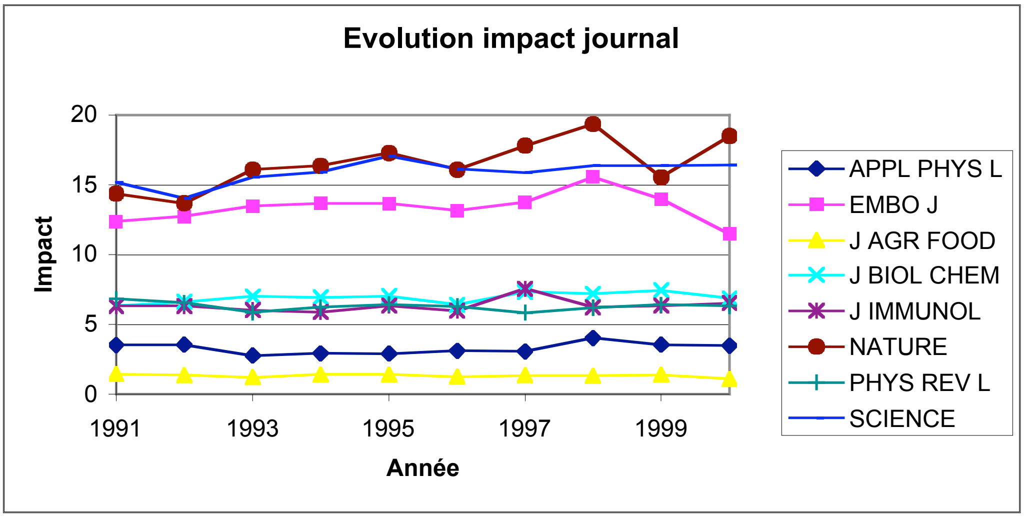
Relying on the analysis of citation patterns for each journal, it was therefore possible to calculate the **expected impact** of each publication.

Among several perimeters, it was chosen to calculate the impact corresponding to all types of references (academic articles but also proceedings, letters...).

3.3.2 Calculating journals' impacts

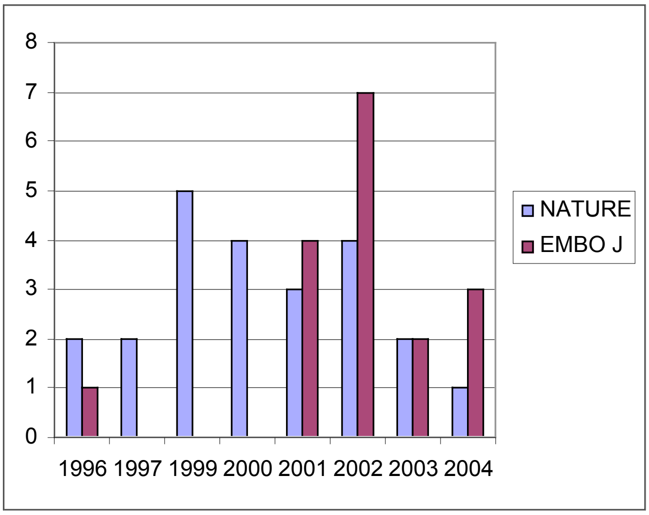
Exhibit 12 shows the time evolution of the impact of the most recurrent journals between 1991 and 2000 (journal impact is not available yet for more recent years).

Exhibit 12 Journals' impact evolution (average number of citations per article, 1991 – 2000)



One observes that journal impacts are rather stable over time, except for two of them: Nature and the EMBO Journal. These variations can appear as problematic since they happen in the years 1998 to 2000, during which most of the articles we analyse were published (see Exhibit 13).

Exhibit 13 Yearly number of articles from the set, published in Nature and the EMBO Journal (1996 – 2004)



In order to take account of these fluctuations, an average journal impact was calculated for the period 1998-1999-2000. This proxy has the advantage of remaining rather simple: considering the imperfect reliability of the whole data set, a complex calculation method would not have been very sensible. In addition, we remind that 94% of the analysed articles were published in 1998 or later. This average can also be considered as a good proxy for the years 2001 to 2004.

The following exhibits allow to compare this average impact to the actual yearly impact values for the most recurrent journals.

Exhibit 14 Comparison between yearly impact values and the 98-00 average for J. Biol. Chem.

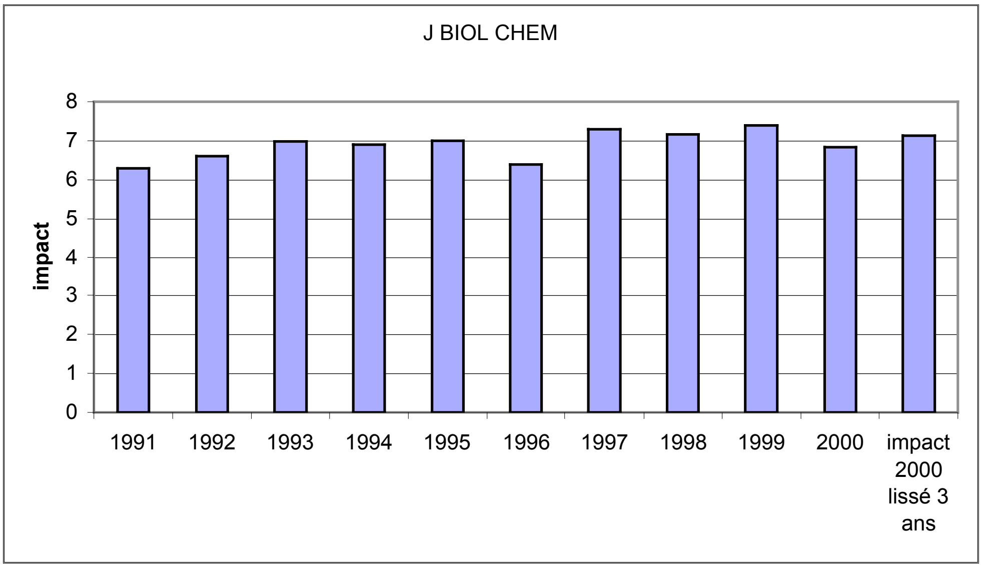


Exhibit 15 Comparison between yearly impact values and the 98-00 average for Nature

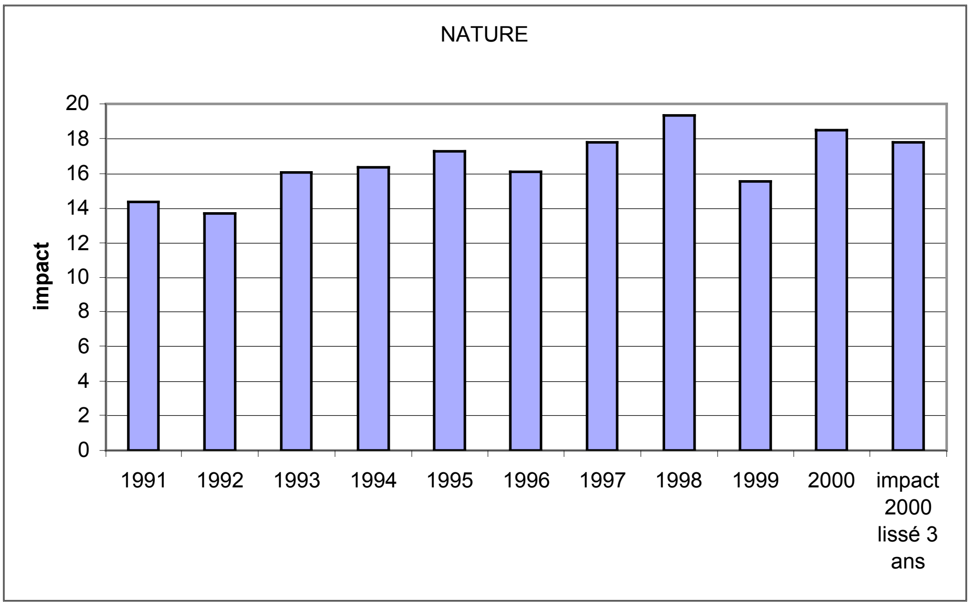


Exhibit 16 Comparison between yearly impact values and the 98-00 average for the EMBO Journal

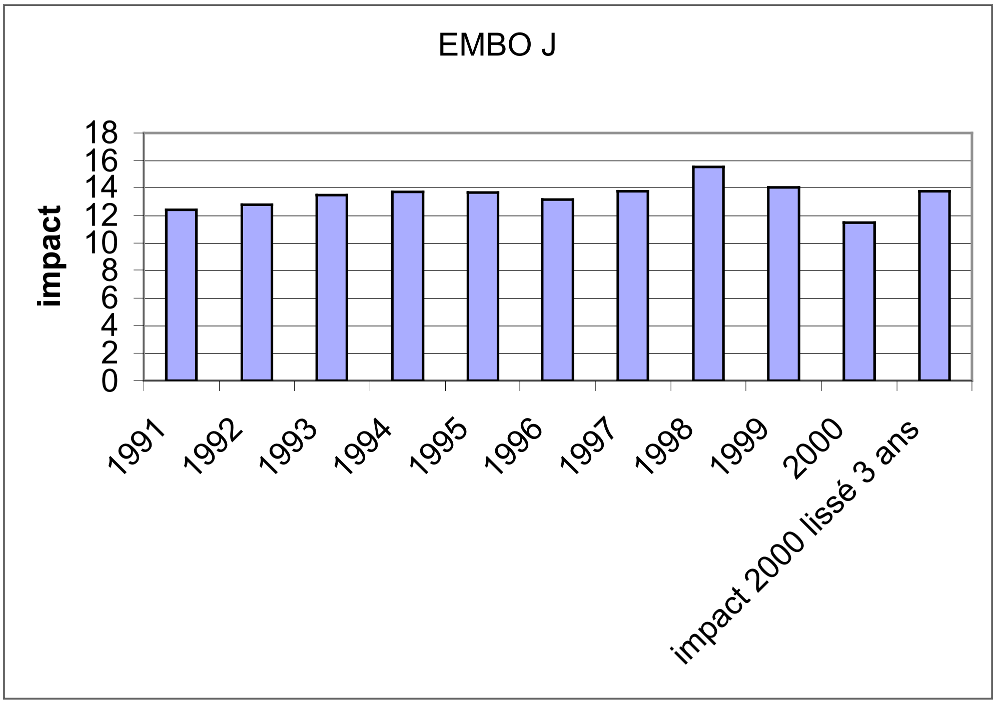
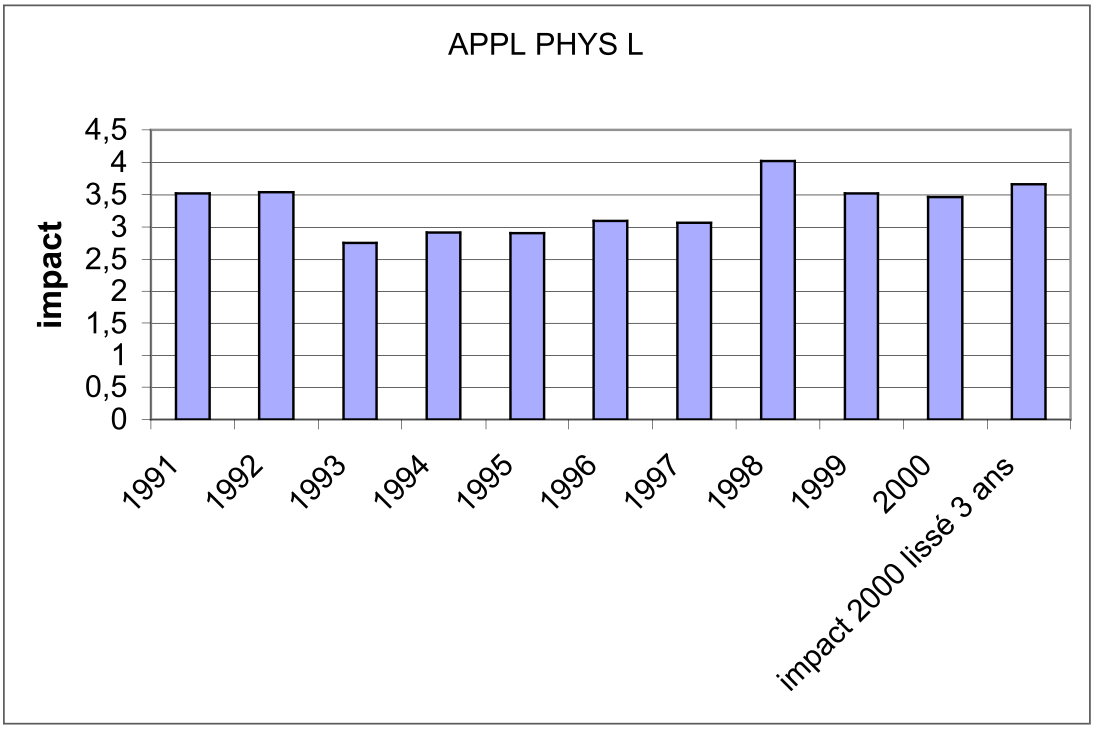


Exhibit 17 Comparison between yearly impact values and the 98-00 average for the Appl. Phys. L.



3.4 Calculating the impact of programmes

Once an average impact has been calculated, it is possible to deduce the estimated impact of each programme from it.

By definition, the programme impact must correspond to **the estimated number of citations per article corresponding to the considered programme**.

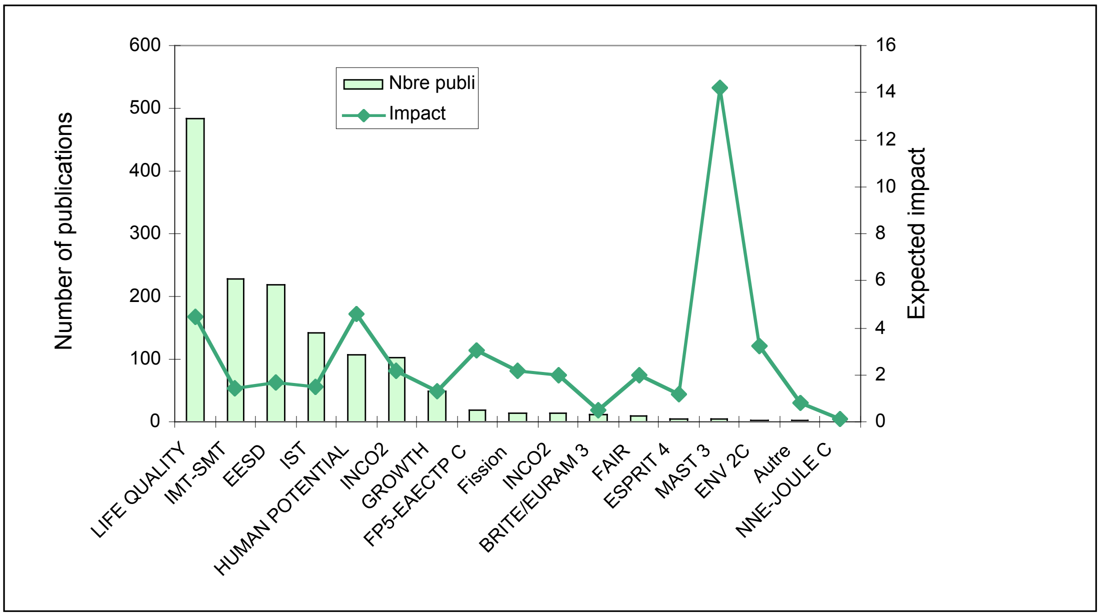
$$\text{programme impact} = \frac{\text{total citations in articles corresponding to the programme}}{\text{number of articles corresponding to the programme}}$$

It is therefore assessed with the following ratio:

$$\text{programme impact} \approx \frac{\sum \text{(journal impact * number of journal appearances in the programme)}}{\text{number of articles corresponding to the programme}}$$

Results for each programme are given in Exhibit 18.

Exhibit 18 Number of corresponding publications and expected impact of each programme



These impacts are generally situated between 1.5 and 4 (apart from MAST 3 whose impact is arguably non significant given the very low number of corresponding publications, 3 out 4 of which are published in Nature). If one only considers programmes with a high number of publications, two of them appear to have relatively higher an impact compared to the others: Human Potential and Quality of Life.

3.5 Calculating the disciplinary scope of programme outputs

3.5.1 Approach

As indicated in previous chapters, each publication has been associated with a journal represented in the ISI database. For the purpose of their core activities, the OST regularly update this database by associating each journal with one or more among ten scientific disciplines.[11] Following a fractional method, a given journal is associated with all relevant disciplines via an adequate coefficient, so that the journal total makes 1:

- a journal mostly involved in a single discipline is associated to that discipline with a coefficient of 1 (e.g. see the publication referenced 'a1122' in Exhibit 19).
- a journal involved in n disciplines is associated to all of these with a coefficient of 1/n (e.g. see the publication referenced 'a2295' in Exhibit 19).

Exhibit 19 Example of disciplinary OST upgrading on the sample of publications emanating from the programme BRITE / EURAM 3

JOURNAL	ANNEE	PROGRAMME	DISCIPLINE	PROPORTION	REF_PUB
ADV COM LET	2002	BRITE/EURAM 3	4	1	a1122
ADV COM LET	2003	BRITE/EURAM 3	4	1	a0759
MAT SCI E A	2003	BRITE/EURAM 3	4	1	a1828
NDT E INT	2003	BRITE/EURAM 3	4	1	a2953
IEEE APPL S	2003	BRITE/EURAM 3	5	0,5	a2295
J PHYS IV	2001	BRITE/EURAM 3	5	1	a2775
J PHYS IV	2001	BRITE/EURAM 3	5	1	a2654
J PHYS IV	2000	BRITE/EURAM 3	5	1	a2775
J PHYS IV	1999	BRITE/EURAM 3	5	1	a2298
J PHYS IV	1999	BRITE/EURAM 3	5	1	a2126
PHYSICA C	2000	BRITE/EURAM 3	5	1	a1089
PHYSICA C	2002	BRITE/EURAM 3	5	1	a1088
IEEE APPL S	2003	BRITE/EURAM 3	7	0,5	a2295

Note: the column "REF_PUB" shows unique identifiers assigned to each publications

The next step consists, for each research programme, in summing disciplinary coefficients of respective disciplines and, then, in calculating the respective proportion of disciplines in the programme outputs.[12]

In the example given in the Exhibit above, the represented disciplines in BRITE/EURAM outputs are *Chemistry* (discipline number 4, with a proportion of 4/12), *Physics* (discipline number 5, with a proportion of 7.5/12) and *Engineering* (discipline number 7, with a proportion of 0.5/12). No other disciplines are represented by this programme's academic outputs.

[11] These disciplinary fields are: Fundamental biology, Medical research, Applied biology and ecology, Chemistry, Physics, Earth and space sciences, Engineering, Mathematics, Multidisciplinary research, Others.

[12] By construction, the total of disciplinary coefficients for a given programme equals the number of publications associated with this programme.

3.5.2 Results

The results of this approach are provided in the present section: disciplinary scope of the whole publication set as well as that of the 6 RTD programmes accounting for more than 100 publications are presented in the following exhibits.

Exhibit 20 Disciplinary scope of entire publication set (OST nomenclature)

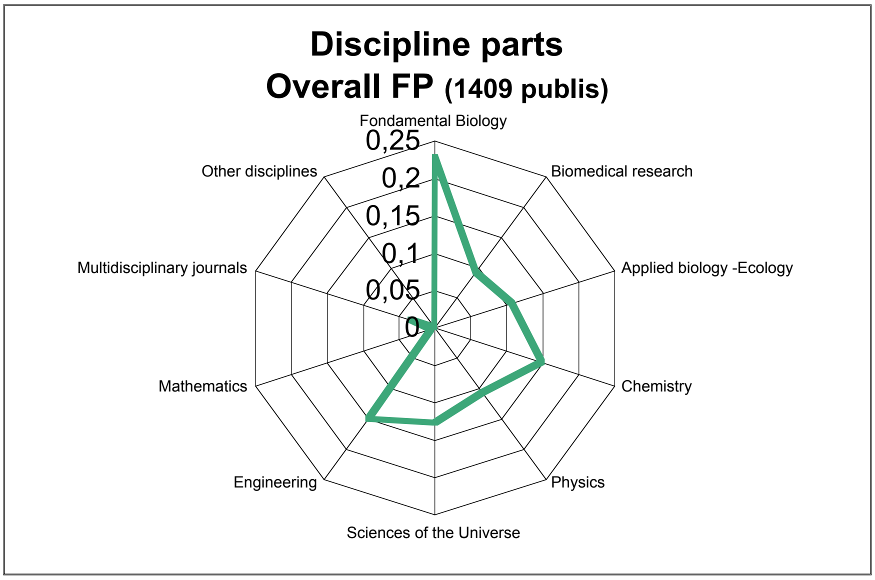
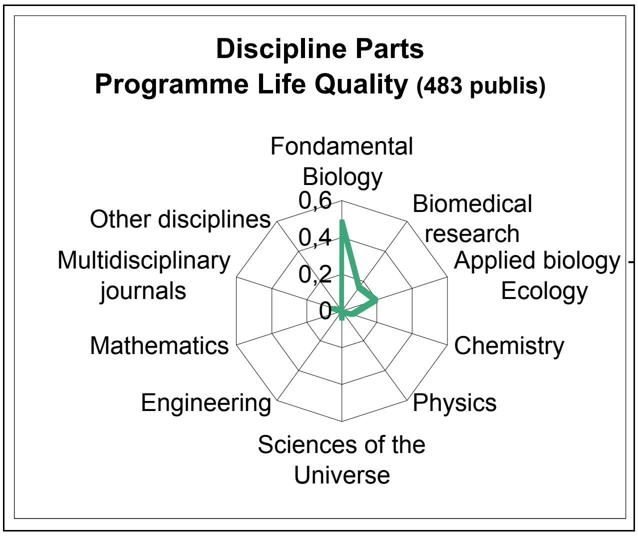


Exhibit 20 illustrates that life sciences (and notably fundamental biology) are the most represented of all fields, beyond matter sciences. When it comes to specific programmes, the following exhibits show the strong thematic focus revealed by the produced publications.

Exhibit 21 Disciplinary scope of publications sets emanating from the programmes Quality of Life and IMT-SMT (according to the OST nomenclature)



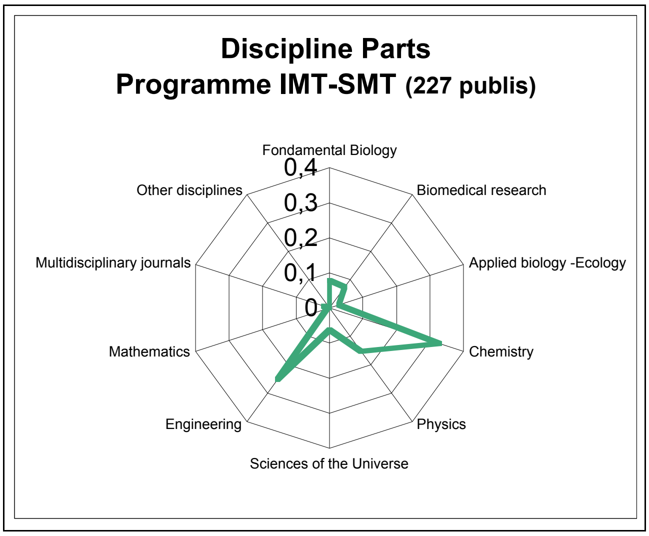
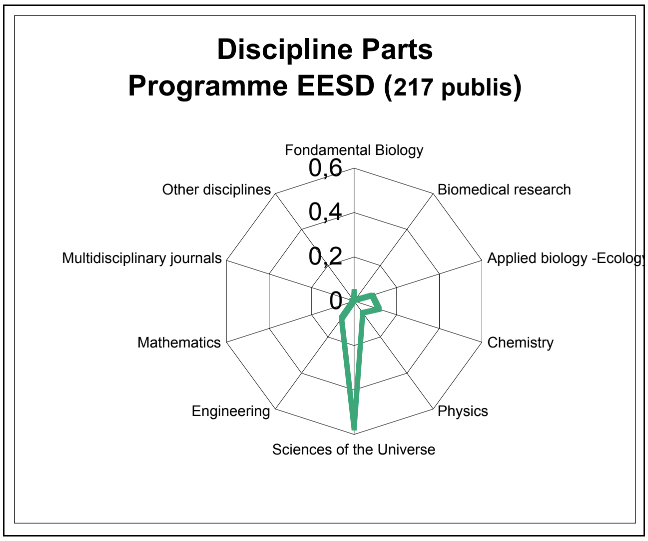


Exhibit 22 Disciplinary scope of publications sets emanating from the programmes EESD and IST (according to the OST nomenclature)



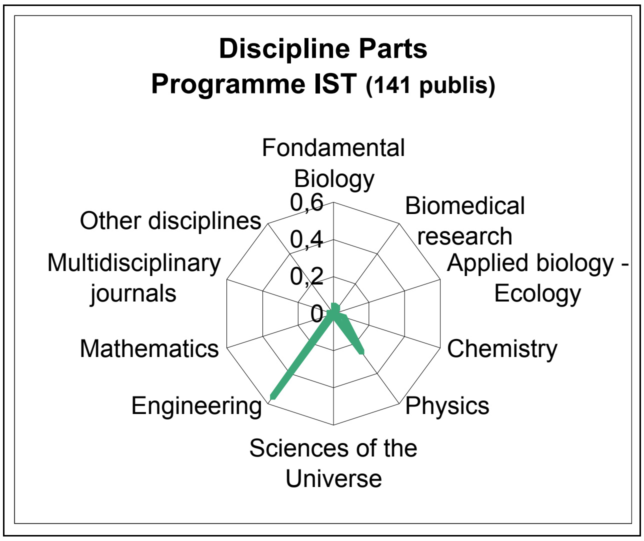
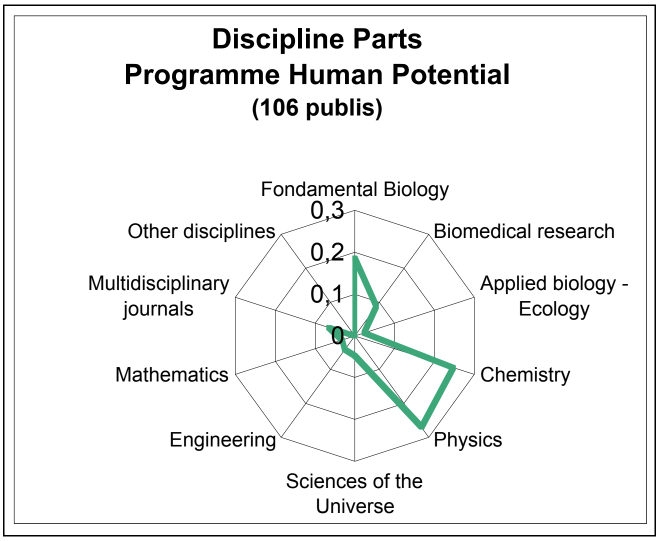
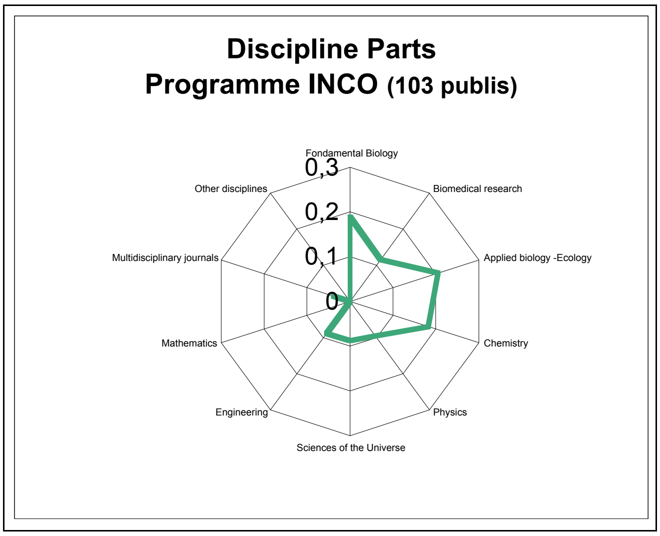


Exhibit 23 Disciplinary scope of publications sets emanating from the programmes Human Potential and INCO (according to the OST nomenclature)





The two past exhibits are interesting in that they provide some insight on the disciplinary scope of multi-disciplinary programmes. Publications emanating from Human Potential appear as mostly focused on chemistry and physics; whereas those coming from INCO are more evenly distributed among disciplines.

4 Complementary analyses

Not only does the previous provide some new and interesting – though fragile – results about the FP impact, but it also demonstrates that the envisaged methodology is feasible, provided the reliability of input data can be improved.

In order to overcome some of the methodological problems that were identified, we carried out 3 kinds of complementary investigations, the results of which are detailed below.

4.1 Assessing the representativeness of data

Although it seems impossible to get a sound estimate of the number of scientific publications which were enabled by EU funded research, it is clear that this number is far more important than 1,410. It is therefore important to know the share of the publications analysed in this study as compared to the total potential number of publications produced with the help of the Framework Programme.

The following exhibit compares, on a virtually exhaustive sample of RTD programmes, the number of projects to which one or more publications could be linked and the total number of projects included in the Cordis database for the main RTD programmes.

Exhibit 24 projects covered by this study compared to total projects of RTD programmes (on a sample of RTD programmes)

RTD programme	# projects (Cordis)	# projects (present study)	Ratio (A)	# publications (B)	Expected total number of publications (B/A)
QoL (5 FWP)	2,321	189	8.1%	483	5,931
IST (5 FWP)	2,375	67	2.8%	141	4,998
GROWTH (5 FWP)	1,970	24	1.2%	50	410
EESD (5 FWP)	1,756	82	4.7%	218	4,668
INCO 2 (5 FWP)	1,026	6	0.6%	13	2,223
HUMAN POTENTIAL (5 FWP)	457	61	13.3%	106	794
EURATOM (5 FWP)	1,285	7	0.5%	19	3,488
BRITE/EURAM 3 (4 FWP)	2,247	2	0.1%	12	13,482
SMT (4 FWP)	426	128	30%	227	755
Sample total	**12,090**	**566**	**4,7%**	**1,269**	**36,751**

Source: CORDIS

It is clear that the publications on which the indicators proposed in the present report were built only come from a very low fraction of EU-funded projects (in the order of magnitude of 5%). The theoretical value of the total number of publications emanating from EU funded programmes would be in the order of magnitude of 40,000, that is 20 to 30 times as much as in the sample in our possession.

Even admitting that not all EU funded projects lead to publications, and that consequently the Commission databases are likely to account for slightly more than 5% of the projects they fund, there is undoubtedly a huge gap between the actual academic impact of EU funded research and the one observed in this study.

4.2 Approaching the effective impact of RTD programmes

As explained in the above sections, the rather poor quality of data does not allow to identify the 1,410 publications from our sample in the SCI database and therefore to calculate the actual impact factor of these publications.

In the perspective of progressing towards the measurement of EU programmes effective impact, the Commission asked us at the end of the year 2004 whether it was possible to identify the authors, which could leave open the opportunity to analyse the dynamic evolution of the impact factor of their publications, with a possible shift in the curve at the time of their participation in an FP project.

The Commission processed the corresponding data in order to extract the names of publications' authors in a separate field and forwarded the file to EPEC for further treatments. The feasibility study was carried out in January 2005 on a sample of three programmes: Quality of Life, IMT-SMT and Brite Euram 3.

Unfortunately, the result of this analysis is negative: homonymic authors are so frequent that this only criterion is not enough to sort all publications included in the SCI database according to their author. In other words, it is not possible to establish sound individual "track records" with SCI-registered publications. In many cases, all publications known by the SCI as published by a given author' spelling (e.g. BROWN D or LI S) are so numerous (we obtained more than 4,000 publications for 156 authors between the years 1991 and 2000 – hence an average of 3.2 publication per year for each of these authors over this whole period) and, especially, connected to so many different disciplinary fields that it must be concluded that these figures represent more than one person.

From a methodological perspective, it is also an additional confirmation that author's name in itself is not the most relevant criterion when it comes to carry out micro-bibliometric analyses (i.e. to identify publications and not only journals). If the Commission is willing to proceed to any further attempt in this way, it is first necessary to extract other information from notices: journal number, pages and publication title.

4.3 Appraising the 'quality' of publications

The third attempt at measuring the impact of EU funded research on a more accurate basis consisted in comparing the average impact factor of the journals in which FWP participants issued their publications to that of all journals from the same discipline. This complementary analysis was made for 3 programmes: Quality of Life, EESD and IMT-SMT. Results are provided below.

The following table (Exhibit 25) shows the average impact factor of SCI-registered journals of each disciplinary field.

Exhibit 25 Average impact factor of SCI-registered journals

Disciplinary field	Journals' average impact factor
Applied biology – ecology	1.01
Fundamental biology	3.18
Chemistry	1.45
Multidisciplinary journals	7.00
Mathematics	0.40
Physics	1.78
Medical research	1.69
Earth and space sciences	1.59
Engineering	0.46
All	1.75

Source: SCI / OST

The following exhibit now displays the comparison between the estimated impact factor of the three RTD programmes, as calculated in the previous section, and the average impact factor of the corresponding disciplinary fields according to SCI-registered journals.

Exhibit 26 RTD programmes estimated IF compared to SCI disciplinary average IF

RTD programmes	Estimated IF (as calculated in previous section)	Average disciplinary IF of the most represented field(s)	
Quality of Life	4.46	3.18 (fundamental biology)	
EESD	1.70	1.59 (earth and space sciences)	
IMT-SMT	1.42	1.45 (chemistry)	1.78 (physics)

These results show some differences among programmes:

- the estimated IF of publications emanating from QUALITY OF LIFE is significantly higher than the SCI average disciplinary IF in the main field of this programme (fundamental biology); in other terms, this provides some evidence that publications issued via this EU funded programme were accepted in a significant proportion by the top journals of the corresponding discipline
- publications issued from EESD projects also show such a trend to be published by top journals, although the gap between "EU science" and "average science" is quite less important
- publications issued from IMT-SMT projects seem to have a lower expected IF than the average publications of the corresponding fields.

5 Conclusion and recommendations

The results of this study demonstrate the validity of the present methodological approach. Even though the dataset is incomplete, **the findings are very promising and this should encourage to undertake further investigations**. This section proposes some concrete ideas to proceed towards more detailed indicators.

There are three types of recommendations to make, relating to respectively:

- Data format
- Data collection procedures
- Other possible indicators

These will be discussed in turn, below.

5.1 Data format: standardise data entry

The preliminary work needed to prepare data for use (Chapter 2), starting from data collected by the Commission, lead to the following recommendations.

First, whichever data collecting process should be retained, it is important to make sure that the computed data do not suffer from any bias, neither in terms of time coverage nor in terms of programme representativeness. Generally speaking, efforts should be directed towards improving the coverage of EU funded projects.

Second, it is of utmost importance to improve the **format** of collected data. All items identified as relevant should be standardised[13] (e.g. multiple choices questions are preferable to free text) and separated in specific fields: journal title, journal number, publication title, publication year, name of author(s)… must therefore not be documented in a single free-text field but appear in separate fields.

It should be noted that such a standardisation will not only improve processing time. It would especially allow to process *more* types of data and therefore to obtain more accurate indicators. The indicators given in the current study rely on the identification *of journal titles only* which explains why only an estimation of *programme* impacts could be given. If publications could have been systematically identified *individually*, that is, if their title, author and publication details had been registered according to standard and separate formats, the following additional analyses could have been easily made:

- Calculation of the actual impact of *publications* emanating from RTD programmes (and not only an average estimation per journal)
- Taking account of *all* publications, including those which had to be removed because of non complete references
- Comparing their impact to that of other publications published in the same journals or countries.

[13] As emphasised by De Laat & Martin, 2003 (op.cit.), in their analysis of TIPs.

Such an improvement implies the standardised collection of:

- the type of publication: articles, notes, letters…
- the title of the journal
- the publication year
- authors
- as well as the detailed references of the EC project which lead to the publication

5.2 Data collection: toward shared indicator platforms

The poor quality of data in the databases used for this study does not so much depend on practical problems encountered by DG RESEARCH but on structural shortcomings. In other words, it is quite unlikely that more numerous, better, and cleaner, data could be obtained via voluntary consultations of project participants. Even if some arrangements could help in standardising the raw results, or obtaining a better response rate, the question of time and programme coverage will remain largely unsolved.

This idea tends to converge with past evaluation of TIPs which found that a systematic and complex procedure is not appropriate to establish a fruitful contact with former participants in EC programmes.

Another, better, approach can be envisaged on the basis of the experience of the OST in building up **shared indicator platforms**:

- DG RESEARCH should involve European research organisations very closely in the process, on a co-operative basis. A proposal would be that these organisations provide the list of all publications emanating from their research labs that can be attributed to some EC project.
- In return, the Commission could quite easily process the data and calculate accurate indicators. In exchange for their voluntary participation, each partner organisation could be provided with global (public?) indicators as well as specific (confidential?) indicators focusing on their respective institutional perimeter.[14]

This would lead to a win-win situation between stakeholders and the European Commission, that would allow DG RESEARCH to produce sound bibliometric impact indicators concerning EC research programmes on a regular basis, and the European research organisations to better position themselves.

5.3 Another possible indicator: patents

The present study was to focus on scientific bibliometric indicators only. However, the analysis of the contents of the two database provided by the Commission made it possible to analyse other interesting options for impact assessment. Patents turn out to be often present as a result of European projects in these two databases. For several reasons, they are probably the only other type of output which could justify a similar effort to establish indicators.

---

[14] This is how, on a "micro"-scale, the S&T Observatory for Angers Technopole, set up by Technopolis France, works: all laboratories of this region cooperate to the endeavour since in return they receive a "positioning" assessment that they can use in their own strategy process.

Like academic publications, patents have the advantage of being internationally classified in publicly (against payment) accessible databases. It is therefore possible – at least theoretically – to envisage a formal comparison between patents emanating from EC research and other sets of patents. Exhibit 27, established within the framework of the present study, indicates the number of intellectual property agreements which were declared to stem from EC programmes (databases as in section 2.1).

Exhibit 27 Number of IP agreements established in the various EC programmes

Programme acronym	No of projects	patents applied for	patents search	patents granted	register ed design	tradema rk appl.	copyrig hts	secret	other
BRITE/EURAM 3	55	15	16	2	2	0	5	19	3
EESD	1338	62	56	20	6	14	230	364	159
ENV 2C	5	0	0	0	0	0	2	1	0
ESPRIT 4	12	2	0	1	0	4	3	2	3
FAIR	65	5	0	1	0	0	2	8	15
FP5-EAECTP C	101	2	2	2	0	0	7	6	13
GROWTH	1436	101	117	35	5	44	159	567	130
HUM. POTENTIAL	41	2	0	0	0	0	3	2	12
INCO	24	2	1	0	1	1	1	5	5
INCO 2	133	7	8	3	5	3	25	20	21
INNOVATION	6	2	1	0	2	2	0	1	0
INNOVATION-SME	31	0	0	0	0	1	10	1	0
IST	2436	141	77	44	20	82	831	504	240
QUALITY OF LIFE	869	91	48	17	7	11	80	163	131
MAST 3	8	0	0	0	0	0	1	0	1
NNE-JOULE C	6	1	1	1	0	0	0	2	0
SMT	1	0	0	0	0	0	0	1	0
TOTAL	**6567**	**433**	**327**	**126**	**48**	**162**	**1359**	**1666**	**733**

For other outputs of the Framework Programmes –software, web pages, policy reports, etc. – now international reference databases exist. Such outputs can only be counted in absolute terms and not compared to any international benchmark.

Annexe A Quick literature review on bibliometrics and impact assessment

This annexe presents the results in reviewing the academic literature about the use of bibliometric tools in the frame of RTD programmes impact assessment.

A first finding is that only a limited number of recent publications on the topic. After a rise of interest for quantitative methods of science impact assessment in the 1990's, and whereas academic literature is still abundant on many other areas of bibliometrics (patent citation analysis, research front analysis…), programme evaluation does not appear as a topical field of application for bibliometric tools.

A.1 Quick overview of main bibliometric indicators[15]

Difficulties in measuring outputs of research projects are known to be numerous: they include at least the existence of uncertainties, the multiplicity of consequences, the cumulative nature of research, and the effect of results' transferability (Melkers, 1993).

It is therefore needed to resort to proxies when one wishes to assess the impact of a given RTD project or programme. Bibliometrics are one of these possible proxies. This section gives a brief overview of the different bibliometric indicators that exist.

Working hypotheses

The implicit assumption lying behind the construction of bibliometric indicators is that they provide a faithful picture of research activity – or at least of some significant part of it. For instance, number and type of publications are considered to be indications of scientific production. Citation-based indicators are regarded as measures of impact or international visibility of research (Narin, 1976; Garfield, 1979; Martin and Irvine, 1983; Moed et al. 1985). While publication data is used to estimate basic research outputs, patent data is typically used to measure "inventiveness, innovation, or technological change" (Papadakis) or technological development (Holbrook).

Still, several limits must be mentioned as to the validity of the above assumptions.

First, as Martin puts it, output indicators are linked with scientific **production** (defined as "the extent to which this consumption of resources creates a body of scientific results embodied both in research publications and in other types of less formal communication between scientists"), but their relationship to scientific **progress** is less direct ("the extent to which scientific activity results in substantive contributions to scientific knowledge"). Indicators of scientific progress are however the critical measure for assessing the results of basic research (Martin, 1996).

Second, as Pavitt suggested very early on (Pavitt, 1984), no one indicator is capable of measuring all aspects of R&D; this continues to be confirmed (Martin, 1996; Melkers, 1993). Therefore, if one remains within the realm of bibliometrics, at least

[15] Cf. Winthrop and al.

different *partial* output indicators must be used in combination in order to obtain a more complete picture of R&D performance, among which the five following:

- publication counts
- citations from journals articles to journals articles
- patent counts
- citations from patents to patents
- citations from patents to journals

Publications counts
Publication counts, the most basic of bibliometric measures, is best used to measure total research output but cannot discern the quality of these outputs (Melkers, 1993). One problem Martin notes with publication is that publications are good measures of scientific production, but inadequate indicators of scientific progress (eventhough academic journals rely on quite severe filtering processes). Most publications make only a small contribution to scientific knowledge, whereas a few seminal pieces make large contributions. A publications count indicator is unable to measure the quality of publication (Martin, 1996).

Another problem noted by Martin is that publication counts reflect not only the level of scientific progress made by an individual or group, but also reflect institutional publication practices, the country of origin, research area, and emphasis on publications for obtaining research funds. Unfortunately, the variance in the publication counts due to effects other than scientific progress cannot be ignored. It is incorrect to assume the effect of scientific progress on publication counts is far greater than the effects of publication practices, country of origin, and so forth. There is also nothing to indicate the total effect of these other influences is random. Over large aggregations or periods of time, the effects do not balance out (Martin, 1996). Another problem with publication counts is each publication represents a different contribution to science. Some authors publish many papers, each representing a small contribution to science while other authors publish only a few papers representing large contributions (Okubo, 1997).

Citations to journal articles from journal articles
Melkers states that citation counts address questions of quality, influence, and transfer of knowledge. It is assumed that the most cited works contribute the most to science (Melkers, 1993). The purpose of citation analysis is to serve as a proxy measure for the contributions to scientific progress (Martin, 1996). “It is presumed that a paper must have a certain quality in order to have an impact on the scientific community” (Okubo, 1997). Some problems with citations include the count of critiques as independent works, failures to cite early works, variations of citation rates across fields and papers, and citations to an author’s own work (Martin, 1996). Other problems include highly cited elementary works (i.e. elementary statistical textbooks) and popular science, which may not be critical science (Lindsey, 1989).

Patent counts
Patent counts, similar to publication counts, are the number of patents produced by an organisation. Papadakis states a patent is a minimum standard of “inventive significance” and represents a base unit. Counts of patent data are counts of patents categorized by firm, type of industry, patent class, nationality of inventor, or other

category (Papadakis, 1993). At least two problems exist with patent data. For various reasons, not all work receives a patent (Papadakis, 1993). Some patents contribute more to technology than others do (Holbrook, 1992).

Citations to patents from patents

Narin and Olivastro note citations to patents from patents are usually the references cited to US patents on the front page of a patent package and are typically the basis for citation analysis (Narin and Olivastro, 1988). Since patent examiners, not the inventors, write these citations, questions arise as to the validity and the completeness of the citations (Okubo, 1997). Narin and Olivastro observe it is assumed that highly cited patents are important advances in technology. Further, most patents are rarely, if ever, cited, with very few patents being cited as many as five times (Narin and Olivastro, 1988).

Citations to journal articles from patents

Besides referencing other patents, patents also reference journal articles. Counting the number of references to journal articles is a means of linking science to technology (Narin and Olivastro, 1988; Okubo, 1997). Although patent applicants need to link their inventions to scientific literature to more easily obtain patents, they also often wish to conceal the essentials of their invention, which may disqualify this metric as a good indicator of scientific and technological advancement (Okubo, 1997).

A.2 What does literature tell us about using bibliometrics for programme evaluation?

We found very few academic papers dealing specifically with the use of bibliometric indicators for programme evaluation or impact assessment[16]. This chapter gives a detailed description of two among the most relevant case studies.

Assessment of US R&D expenditures' impact on aerospace industry[17]

Winthrop et al. use a multivariate analysis to assess the impact of some governmental R&D expenditures. While they acknowledge that "correlation is not causation", their case study shows a strong relation between government R&D expenditures for the USAF and the aerospace industry, either in terms of publications, patents and GDP. A strong relationship is also found between DoD/NASA R&D expenditures and the aerospace industry.

According to the authors, this relationship between the proxy measures of scientific output and federal government R&D expenditures suggests that observed decreases in federal funding could have a marked impact on national technological advancement.

Measurement of Japanese R&D programmes' additionality

(Hayashi, 2003) recalls that, to justify public intervention in research activities – especially those conducted by private companies, RTD programmes are expected to have an effect which would not have been possible without public intervention. He

[16] This scarcity of past examples remains true even if one does not restrict oneself to academic literature: as an instance, google indicates 32 hits only for the following query ("impact assessment" + "research programme" + bibliometrics") while it indicates more than 16 000 hits for ("impact assessment" + "research programme").

[17] This paper comes from a study partially financed by the Air Force National Laboratory.

thus defines the notion of additionality as the measurement of the change induced by a given public programme in the participants' R&D activities.[18] Considering that, in practice, RTD projects cannot be conducted by researchers who don't benefit from a certain expertise which relies on past research projects and imply forthcoming new ones, this additionality can only be relative.

He then resorts to **patent** bibliometrics to measure the level of continuity / adaptability of research topics investigated by public programmes' participants. He also investigates the possible similarities between research groups participating in a same project. His conclusions vary from one field to another.

A.3 Bibliometrics compared to other evaluation tools in RTD programme management

Since bibliometric tools were introduced, academic debate on their relevance and accuracy has remained open – as well as the corresponding debates on other evaluation tools. This chapter provides some insights on (still open) discussions about using bibliometrics for RTD programme impact assessment.

Bibliometrics compared to other quantitative indicators

Economic impact assessment in the framework of the US GPRA

In his paper, Tassey gives an overview of economic impact assessment methods used by the US National Institute of Standards and Technology (NIST). He recalls indeed that, since the adoption of the Government Performance and Results Act (GPRA) in 1993 in the USA, both strategic planning and subsequent economic impact assessments must be undertaken to manage all programmes that were approved by the Congress.

He first makes a distinction between technical **outputs** and economic **outcomes**. Outputs are the activities performed and the products produced by NIST research. They describe the nature of the immediate impact of the research on industry. Such outputs are often purely technical, but they can also include institutional effects. On the other hand, outcomes are the bottom-line economic impacts from projects (R&D efficiency, manufacturing productivity, quality enhancement, facilitation of market transactions, etc.). These outcomes are measured quantitatively whenever possible, preferably in one of several generally acceptable forms (see below). According to the author, "some frequently used measures such as publication counts, patents, and citation indices **are not considered particularly useful metrics**". In addition to these statistics, the studies often document anecdotal numerical outcome information such as market share impacts. Finally, qualitative outcomes such as impacts on R&D investment decisions, changes in production and quality strategies, and new market strategies are documented. The author concludes pointing that "these economic impact studies contain a mixture of quantitative and qualitative outcomes", following an approach which provides "a more comprehensive and balanced view of a projects impacts than a more focused study in which a single, end-point quantitative impact measure is the only objective".

[18] For a definition of different types of additionality as used in R&D evaluation see Georghiou…

Tassey also mentions the several quantitative metrics used in measuring projects outcomes and points that they are “consistent with those employed by economists and corporate finance analysts to assess corporate and government R&D projects”:

- Net Present Value (NPV): This is an absolute measure of the value or expected value of an R&D project in constant [euro] terms. It is calculated by discounting (or inflating) the cost and benefit time series to the reference year and subtracting the present value of costs from the present value of benefits to yield the net present value of the investment. Varying forms of this measure are widely used by industry, where it is often referred to as “discounted cash flow”. Like most metrics, its use is affected by the selection of a discount rate which is used to adjust the time series of benefits and costs for risk, time preferences of money, and inflation. This selection is not straightforward because of different views with respect to how many of these three factors to include in determining the discount rate.
- Benefit-Cost Ratio (BCR): Calculated from the same estimates of the present value of benefits and costs as are used for NPV, except that the two values are expressed as a ratio to provide a measure of relative efficiency across projects. As a BCR=1 is the break even value, presumably any project with a greater value is a successful project.
- Social Rate of Return (SRR): Calculated as the discount rate that is needed to reduce the time series of net benefits under analysis to zero (i.e. the break even point). Economists often use this metric (which is called the “internal rate of return” in the corporate finance literature) because, like a BCR, it provides an indicator of relative efficiency. One feature that distinguishes this measure from alternatives, including other rate-of-return measures, is that its value depends solely on the internal characteristics of the project (i.e. selection of a discount rate is not required).
- Adjusted Internal Rate of Return (AIRR): The AIRR is a measure of the annual yield from a project over a given study period, taking into account the reinvestment of interim receipts at a specified rate. When the reinvestment rate is made explicit, all investment costs are easily expressible as a time-equivalent initial outlay and all non investment cash flows as a time-equivalent terminal amount. This allows a straight-forward comparison of the amount of money that comes out of a project (its terminal value) with the amount of money put into a project (its present value of investment costs).

Impact assessment of three South African research programmes in the field of environment

Esterhuizen and Liebenberg give another example of an RTD programme impact assessment which relied on non bibliometric indicators. Their paper reports on the use of indicators within a “comprehensive impact assessment approach” applied to three case studies of the Agricultural Research Council (ARC) of South Africa. Three broad categories of impact assessment form part of this comprehensive approach:

- direct outcome of research activities
- institutional impact
- impact.on the level of competences

Indicators are said essential for measuring impacts on various levels. Two major groups of indicators are defined:

- impact indicators: the most difficult to measure and collect, mainly because of lags between project implementation and impact (or put another way, between the time of impact and the time it is feasible to collect data relating to impact).
- performance indicators: measures of project outcomes, outputs, and inputs that are monitored during project implementation to assess progress toward project objectives.

Bibliometrics compared to peer-review

The argument about respective advantages of qualitative and quantitative methods goes far beyond the limits of this survey (see Mohr 1999 for an example of epistemological justification of qualitative impact analysis). Would one only stick to the debate between bibliometric indicators and peer-review, one would already fall on an abundant literature[19].

Part of these studies have considered the correlation between indicators and peer-review at the stage of the publishing process (Daniel, 1993; Korevaar and Moed, 1996). Others compare bibliometric indicators and peer judgement on the basis of questionnaires among scholars and experts (Anderson et al., 1978; Martin and Irvine, 1981).

In their own study published in 1998, Rinia et al. focus on the correlation between bibliometric indicators and the outcomes of peer judgements made by expert committees of physics in the Netherlands[20]. On the basis of a set of more than 5 000 publications and nearly 50 000 citations, their study shows varying correlations between different bibliometric indicators and the outcomes of a peer evaluation procedure:

- Positive and significant but no perfect correlations are found between some bibliometric indicators and peer judgements of the analysed research programme.
- At the level of overall peer judgements, the highest correlations are found between peer-judgement and the average number of citations per publication (CPP) and the citations averages normalised to world average (CPP/JCSm and CPP/FCSm)[21].
- However, the impact of journal alone, as reflected by the mean journal citation rates, **does not correlate well with the quality of scientific project as perceived by peers**. Authors indeed underlie that the mean citation rate of a given journal or a journal category is only partly explained by its relative impact: it also reflects differences between citation patterns from one field to another. Moreover, the impact of articles published in one given journal may vary strongly. According to the authors, this result then “supports the conclusion that journal impact factor alone does not provide a sufficient basis for assessing research performance”.
- Correlation between bibliometric indicators and peer judgement is higher in the case of exploratory research than in the case of application-driven research.

---

19 Anderson et al., 1978; Bayer and Fulger, 1966; Chang, 1975; Cole and Cole, 1967; Martin and Irvin, 1983; Nederhof, 1988; Nederhof and van Raan, 1987; Nederhof and van Raan, 1989.

20 As a first step, they focused on the results of an evaluation of 56 research programmes in condensed matter physics in the Netherlands, a subfield which accounts for roughly one third of the total of Dutch physics.

21 These two indicators respectively compare the average number of citations per publication with the world standards of the corresponding journals and fields.

- The shorter the period on which bibliometric indicators are calculated, the lower the correlation can be observed with peer-judgement.

A.4 More common uses of bibliometrics

The previous chapter meant to give insights on other possible ways – both qualitative and quantitative – to assess the impact of RTD programmes. The purpose of this chapter is, on the other hand, to give an overview of other possible uses of bibliometrics.

Measuring performance or impact

Bibliometrics are often resorted to when it comes to appraise the impact or the leverage effect of a given body of knowledge. Papers mentioned below offer good examples of such studies which commonly apply to multiple "objects" but research programmes.

First, bibliometrics can be used to appraise the level of academic excellence of articles published by a given journal. It can reflect the journal board's own wish to account for quality (see Knight, Hult and Bashaw for the analysis of the research published in the Journal of Business Research over the period 1985-1999) as well as it can respond to research actors' need to have an accurate measure of the perception of a given journal (see Turban, Zhou and Maa for a presentation of a formal procedure that integrates objective and subjective judgments to provide a comprehensive evaluation method).

Second, in quite the same perspective, one will find several studies seeking to assess the "quality" or "impact" of research for a given country or group of countries and a given research topic. As an instance, Mela, Cimmino et Ugoloini provide in their paper the results of a study into the impact of EU oncology research as compared to other countries. One could label it as a "purely bibliometric" approach, linking citation analysis to the notion of research impact and key words analysis to that of research subfields. Ugolini and Mela then updated their own study in 2003.

Obviously, as soon as one relies on the assumption that the impact factor is a good proxy for research quality, it is very tempting to use it as a key indicator to rank individuals or institutions. The British Research Assessment Exercise provides a well-known instance of bibliometrics contributing to ranking institutions (here universities). In the same order of idea, Baden-Fuller, Ravazzolo and Schweizer present an attempt at ranking research activities of all business schools in Europe, based on work published in top quality international journals. According to the authors, the comparison of this ranking with other ones (such as that of The Financial Times and national ranking schemes) allows to conclude that this list is more robust and comprehensive.

As for individuals, lots of ranking exercises are carried out without justifying any academic paper. More often than not, papers dealing with this subject will try and adjust/upgrade the quantitative tools used for such ranking. For instance, as regards their disciplinary field (educational psychology), Hsieh et al. argument how the previously dominant notion of productivity (having few co-authors and high authorship placement) needs updating and taking account of co-authorship as a indicator of productivity. According to the authors, such updating could change significantly previously published productivity ranking of scientists. Another example lies in Hollis' paper, which uses panel data on 339 economists to evaluate the relationship between co-authorship and output. It is shown that for a given individual, more co-authorship is associated with higher quality, greater length, and greater frequency of publications[22].

Content / corpus analysis
Biliometrics are historically derived from bibliography and documentation management. It is therefore unsurprising to fall on documentation management papers, which explain how bibliometric indicators can contribute to run some advanced management of libraries' set of journals. As an instance, Enssle and Wilde describe how indicators (internal statistics on journal usage merged with statistics from commercial sources) can help in the "cancellation decision process" of the many university libraries exposed to the impact of inflation on journal prices. Whereas deciding which journals should be cancelled is a serious and frustrating task, statistical information about a library collection can be a valuable tool in both identifying journals for cancellation and justifying cancellation decisions.

In addition, bibliometrics used by research analysts may also stick to descriptive analysis and information mapping and avoid normative ranking or quality assessment such as described in previous paragraph. Bibliometric tools thus often serve to spot topical research fronts within a broader discipline. This is a very common kind of study as far as basic research is concerned: Zhang provides an example of such research fronts analysis concerning medical research in Japan over the 1990-1992 period. Waugh and Ruppel also report on some citation analysis (concerning the specific theme of Workforce Education and Development) which allowed to determine core serials in the discipline, provide the university library with a guide to serials acquisition and maintenance in the discipline, and provide future students with a core list of journals.

Still, another interesting but less common example is given by van Raan and van Leeuwen, who present the results of a study into topical research fronts and scientific basis of a "typically interdisciplinary, applied field" (nutrition and food research). They claim to have based their approach on "an advanced bibliometric analysis with novel elements to assess the influence and dissemination of research results and to measure interdisciplinarity". According to the authors, this case study shows that bibliometric allows assessment beyond conventional academic standards.

Linkages between science and innovation and patent bibliometrics
This paragraph describes a third major type of bibliometric study: the analysis of knowledge flows between actors of research and those of innovation. This topical question goes far beyond the boundaries of bibliometrics since, as Tijssen writes it, "there is still an urgent need for comprehensive models, reliable data and analytical tools to describe and monitor links between R&D and industrial innovation in more

[22] However, the net relationship between co-authorship and output attributable to the individual is negative after discounting for the number of authors. The results of this paper suggest that universities and granting agencies which preferentially reward research collaboration may be undermining their goal of maximizing research output.

detail", eventhough several innovation studies and surveys have provided some convincing empirical evidence of impacts and benefits of research to technical progress.

Some authors thus put forward that counting patents related to basic research activities – as far as one could judge from patent citations – would be one way of measuring the relevance of basic research to industry (Narin et al., 1995). Their results indicate a growing relationship between science and technology in a very general way. Still, as Meyer puts it, "this idea of an increasingly science-based technology might convey the impression that there is direct knowledge-transfer taking place" and being reflected in citations to scientific research papers in patents. Still, on the basis of a study of front pages of patents in the field of nanoscale technologies, Martin argues that citation linkages hardly represent a direct link between cited paper and citing patent. Tijssen also presents a different approach, based on a nation-wide mail survey amongst inventors working in the corporate sector and the public research sector in The Netherlands. He also concludes that "citations in patents referring to basic research literature are found to be invalid indictors of a technology's science dependence".

Apart from S&T analysts, investors form another public likely to analyse citation links between publications and patents since they wish to know how far the ownership of patents is a reliable indicator of a given company's forthcoming commercial success. As far as SME are concerned, Hirschey and Richardson state that "scientific measures of patent quality" give investors a useful basis upon which to judge the economic merit of the firm's inventive and innovative activity. Especially in the case of small cap and relatively low price-earning high tech companies, they find "a favourable stock-price influence when both the number of patents, the scientific merit of those patents, and R&D spending is high".

Lastly we only mention here the even vaster field of patent bibliometrics, that is bibliometrics applying to patents only and not to citation links between patents and publications. Among many applications, this subfield of bibliometrics can contribute to the study of knowledge flows between science and technology. Yamin and Otto thus examine the influence of inter-and intra organisational knowledge flows on innovative performance in multinational enterprises, taking the biopharmaceutical sector as a field for application. George, Zahra and Wood partly resort to patent bibliometrics in order to analyse the impact of academic alliances undertaken by biotechnology firms[23].

Studies on bibliometric indicators

The fourth and last subfield of bibliometrics we would mention in this short survey of academic literature is what one could name "pure bibliometrics", i.e. some statistical approaches of bibliometric tools. Here again, this domain is too vast for our survey to give a detailed description of it. Following references are provided as a representative

[23] They show that companies with university linkages have lower R&D expenses while having higher levels of innovative output . However, the results do not support the proposition that companies with university linkages achieve higher financial performance than similar firms without such linkages.

sample of the multiple subjects that can be addressed by the community of bibliometricians.[24]

A.5 References

Anderson R. C., Narin F., Mc Allister P. Publication ratings versus peer ratings of universities. J. Am. Soc. Info. Sci. 29 (1978), 91-103.

Baden-Fuller Charles, Ravazzolo Fabiola and Schweizer Tanja. Making and Measuring Reputations. Long Range Planning 33 (2000) 621-650.

Balconi Margherita, Breschi Stefano, Lissoni Francesco. Networks of inventors and the role of academia: an exploration of Italian patent data. Research Policy 33 (2004)127–145.

Bayer A. E., Fulger F. J. Some correlates of a citation measure of productivity in science. Socio. Educ. 339 (1966), 381-390.

Beile Penny M., Boote David N. and Killingsworth Elizabeth K. A Microscope or a Mirror? A Question of Study Validity Regarding the Use of Dissertation Citation Analysis for Evaluating Research Collections. The Journal of Academic Librarianship, (article in press).

Chang K. H. Evaluation and survey of a subfield physics: magnetic resonance and relaxation studies in the Netherlands. FOM report n° 37 175, Utrecht.

Cole S., Cole J. R. Scientific output and recognition: a study in the operation of the reward system in science. Am. Sociol. Rev. 32 (1967) 377-390.

Daniel H. D. Guardians of science: fairness and reliability of peer-review. 1993. VCH, Weinheim.

Enssle Halcyon R., Wilde Michelle L. So you have to cancel journals? Statistics that help. Library Collections, Acquisitions, & Technical Services 26 (2002) 259–281

Esterhuizen J. M. C., Liebenberg G. F. The use of indicators within a comprehensive impact assessment approach in the three South African research programmes. Agriculture, Ecosystems and Environment 87 (2001) 233-244.

Garfield E. Citation indexing – its theory and applications in science, technology and humanities. New York: Wiley, 1979.

George Gerard, Zahra Shaker A., Wood D. Robley. The effects of business–university alliances on innovative output and financial performance: a study of publicly traded biotechnology companies. Journal of Business Venturing 17 (2002) 577–609.

Gusmao Regina. Developing and using indicators of multilateral S&T cooperation for policy making: the experience from European research programmes. Scientometrics. Vol. 47, No 3 (2000) 493-514.

[24] See for instance Lu and Junping, for an assessment of the statistical correlation between the journal impact factor (JIF) indicator and web-based impact indicators. Hogenboom, for an analysis of citation patterns to US government information. Tsai, for an investigation on information landscaping, i.e. the integration of information mapping, charting, querying and reporting techniques. Beile, Boote and Killingsworth, for a study on the use of dissertation citation to assess the doctoral students' expertise in their use of the scholarly literature and the relevance of dissertation citation analyses for purposes of developing research collections.

Hayashi Takayuki. Bibliometric analysis on additionality of Japanese R&D programmes. Scientometrics. Vol. 56, No 3 (2003) 301-316.

Hirschey Mark, Richardson Vernon J. Are scientific indicators of patent quality useful to investors? Journal of Empirical Finance 11 (2004) 91–107

Holbrook J.A.. Basic indicators of scientific and technological performance. Science and Public Policy 19 (1992), 267–273.

Hollis Aidan. Co-authorship and the output of academic economists. Labour economics 8 (2001) 503-530.

Hogenboom Karen. Has government information on the Internet affected citation patterns? A case study of population studies journals. Journal of Government Information 29 (2002) 392–401.

Hsieh Pei-Hsuan, Acee Taylor, Chung Wen-Hung, Hsieh Ya-Ping, Kim Hyunjin, Thomas Greg D., You Ji-in, and Robinson Daniel H. An alternate look at educational psychologist's productivity from 1991 to 2002. Contemporary Educational Psychology 29 (2004) 333–343.

Knight Gary A., Hult G. T. M., Bashaw R. Edward. Research Productivity in the Journal of Business Research: 1985–1999. Journal of Business Research 49 (2000), 303–314.

Korevaar J. C., Moed H. F. Validation of bibliometric indicators in the field of mathematics. Scientometrics 37 (1996), 117-130.

Lindsey, D. Using citation counts as a measure of quality in science: measuring what's measurable rather than what's valid. Scientometrics 15 (1989)189–203.

Lowe Alan, Locke Joanne. Perceptions of journal quality and research paradigm: results of a web-based survey of British accounting academics. Accounting, Organizations and Society xxx (2004) xxx–xxx (article in press).

Lu An and Junping Qiu. Research on the Relationships between Chinese Journal Impact Factors and External Web Link Counts and Web Impact Factors. The Journal of Academic Librarianship 30 (2004) No 3, 199–204.

Martin, B.R. The use of multiple indicators in the assessment of basic research. Scientometrics 36 (1996), 343–362.

Martin B. R., Irvine J. Assessing basic research. Some partial indicators of scientific progress in radio astronomy. Research policy 12 (1983) 61-83.

Martin B. R., Irvine J. Internal criteria for scientific choice: an evaluation of research in high-energy physics using electron accelerators. Minerva XIX (1981), 408-432.

Mela G. S., Cimmino M. A., Ugoloini D. Impact assessment of oncology research in the European Union. European Journal of Cancer 35 (1999) 1182-1186.

Meyer Martin, Utecht Jan Timm, Goloubeva Tatiana. Free patent information as a resource for policy analysis. World Patent Information 25 (2003) 223–231

Meyer Martin. Does science push technology? Patents citing scientific literature. Research Policy 29 (2000) 409–434.

Mohr L. B. The qualitative method of impact analysis. American journal of evaluation 20 (1999) 69-84.

Melkers, J. Bibliometrics as a tool for analysis of R&D impacts. In: Bozeman, B., Melkers, J. (Eds.), Evaluating R&D Impacts: Methods and Practice. Boston: Academic Press, 1993.

Moed H. F., Burger W. J. M., Frankfort J. G., van Raan A. F. J. The use of bibliometric data for the measurement of university research performance. Research policy 14 (1985) 131-149.

Narin F., Hamilton K. S., Olivastro D. Linkage between agency supported research and patented industrial technology. Research Evaluation 5 (1995) 183-187.

Narin, F., Olivastro, D. Technology indicators on patents and patent citations. In: Van Raan, A.F.J. (Ed.), Handbook of Initiative Studies of Science and Technology. NewYork: Elsevier, 1988.

Narin F. Evaluative bibliometrics: the use of publication and citation analysis in the evaluation of scientific activity. National Science Foundation, 1976, Washington DC.

Nederhof A. J. The validity and reliability of evaluation of scholarly performance. In: van Raan A. F. J. (Ed.), Handbook of quantitative studies of science and technology. Amsterdam: Elsevier/North-Holland, 1988.

Nederhof A. J., van Raan A. F. J. A validation study of bibliometric indicators: the comparative performance of cum laude doctorates in chemistry. Scientometrics 17 (1989) 427-435.

Nederhof A. J., van Raan A. F. J. Peer-review and bibliometric indicators of scientific performance: a comparison of cum laude doctorates with ordinary doctorates in physics. Scientometrics 11 (1987) 333-350.

Okubo, Y. Bibliometric Indicators and Analysis of Research Systems: Methods and Examples. Organisation for Economic Co-Operation and Development, Paris, 1997.

Papadakis, M. Patents and the evaluation of R&D. In: Bozeman, B., Melkers, J. (Eds.), Evaluating R&D Impacts: Methods and Practice. Boston: Academic Press, 1993.

Pavitt, K. Science and technology indicators: eight conclusions. Science and Public Policy 11 (1984), 21–24.

Pilkington Alan, Dyerson Romano, Tissier Omid. The electric vehicle: Patent data as indicators of technological development. World Patent Information 24 (2002) 5–12.

Rinia E. J., van Leeuwen Th. N., van Vuren H. G., van Raan A. F. J. Comparative analysis of a set of bibliometric indicators and central peer-review criteria. Evaluation of condensed matter physics in the Netherlands. Research Policy 27 (1998) 95-107.

Tassey Gregory. Lessons learned about the methodology of economic impact studies – the NIST experience. Evaluation and Program Planning 22 (1999) 113-119.

Tijssen Robert J.W. Science dependence of technologies: evidence from inventions and their inventors. Research Policy 31 (2002) 509–526.

Tsai Bor-sheng. Information landscaping: information mapping, charting, querying and reporting techniques for total quality knowledge management. Information Processing and Management 39 (2003) 639–664.

Turban Efraim, Zhou Duanning, Maa Jian. A group decision support approach to evaluating journals. Information & Management 42 (2004) 31–44.

Ugolini D., Mela G. S. Oncological research overview in the European Union. A 5-year survey. European Journal of Cancer 39 (2003) 1888–1894.

van Raan A. F. J., van Leeuwen Th. N. Assessment of the scientific basis of interdisciplinary, applied research. Application of bibliometric methods in Nutrition and Food Research. Research Policy 31 (2002) 611–632

Waugh C. Keith and Ruppel Margie. Citation Analysis of Dissertation, Thesis, and Research Paper References in Workforce Education and Development. The Journal of Academic Librarianship 30 (2004) No 4, 276–284.

Winthrop Michael F., Deckro Richard F., Kloeber Jack M. Government R&D expenditures and US technology advancement in the aerospace industry: a case study. Journal of engineering and technology management 19 (2002) 287-305.

Yamin Mo, Otto Juliet. Patterns of knowledge flows and MNE innovative performance. Journal of International Management 10 (2004) 239– 258.

Zhang Haiqi. Analysing the Research Articles Published in Three Periodicals of Medical Librarianship. Intl. Inform. & Libr. Rev. (1995) 27, 237–248.

Annexe B Detailed insight in the line-by-line cleaning process carried out by the OST

From the Excell database delivered by Technopolis France to the OST, only 1410 out of the 2387 available references (59%) could be used as inputs for bibliometric investigations.

This paragraph analyses the various reasons which made the rejected references unsuitable for further treatments. 8 causes were identified:

- **uncomplete reference** (ur): it sometimes happened that the given reference did not allow to assert the type of publication (journal, book, conference…) or that the name of the journal was missing. Lack in the reference was found to be the main reason to reject a reference (331 references were concerned). Examples:
  - 'Madsen P. Kongshøj Munch-Madsen P. and Langhoff-Roos K. (2002): 'Employment policies and Social Exclusion: An analysis of the performance of European Employment Regimes' in Muffels R., Tsakloglou' : missing type of publication, missing name of the publication support
  - 'Benthic copmmunity response to sand dredging aand shhopreface nourishment in Dutch coastal waters 2001 Senckenbergianan maritima' : type of publication is missing
  - Manuscript Teves, F., Casqueiro, J., Naranjo, L., Ullán, R.V., Bañuelos, O., Martín, J.F. (2004). Characterization of the lys3 gene and a mutant allele from Penicillium chrysogenum encoding a bi-functional protein with Homoaconitase activity and regulatory functions : missing type, missing name of publication support
  - research publication 102. De Gennaro V., Delage P., Cui Y-J., Schroeder Ch. & Charlier R."Time-dependent behaviour of oil reservoir chalk : a multiphase approach, Soil and Foundations . Paper No 2800 2003 - (sous presse) : the name of publication support is indicated : 'Soil and Foundation', type of publication missing but the information 'Paper number' could indicate a conference type

- **publications related to conferences** (257 references were concerned). It was not always possible to know whether the publication was published in proceedings or only as a communication on the conference site. Examples:
  - ' Proc. CGI 2003' : proceedings were edited
  - 'Comparison of Spray Simulation and Measurement for a Multi-Hole Injector with and without Wall Impingement D. Kuhnke:, ICLASS 2003, Sorrento' : paper title, author's name, conference name and year are given. No mention of proceeding edition

- **unknown journals**. Some journals presented in the claimed references are absent from the OST-SCI journal database (233 references). Several reasons were identified:
  - The disciplinary field of the journal was sometimes out of the scope of the OST database (for some journal related to economics, archaeology…). Examples:
    - 'Comparisons of observed process quality in early child care and education programs in five countries. Early Childhood Research Quarterly 11 1996 447-476'
    - 'Paper in Geoarcheology A paper was submitted to Geoarcheology in May 2002 by McNeill et al. focussing on the emerged shorelines on the East Eliki fault footwall'
  - Some journals did not appear amongst those considered as top-level by ISI nor representative of the worldwide publications patterns in a given topic. This concerned regional journals (from Baltic or Slave countries as well as journals written in Spanish…). Examples:
    - 'Publication Finnish-Swedish Flame Days, 24-25 September 2002, Vaasa, Finland'

  - ‘Journal paper G Fazekas, M Fehér, G Arz, A Tóth, Gy Stefanik, Zs Boros: New Generation of Medical Aids for Patients with Upper Limb Paresis: Application of Robots in Physical Medicine and Rehabilitation (in Hungarian). Rehabilitáció. 2002;12(3):6-9’
- Some reference to journals were erroneous or ambiguous; some references were also very difficult to identify because of non standard abbreviated names / publication type identifiers. Examples:
  - ‘Publication CR Geosciences, 2002’ : should it be matched with GEOSCIENCE CANADA (from the SCI-OST database) ?
  - ‘PEYRET F. JURASZ J. CARREL A. ZEKRI E. GORHAM B. “The Computer Integrated Road Construction Project” Int. journal Automation in Construction vol. 9 p. 447-461 2000’ : the journal is not ‘International journal of Automation in construction’ but ‘Automation in construction’ : the type of journal (international journal) and journal name (Automation in Construction) are mixed and confusing.
- Some journals were created after 2001, the last updating year of the OST-SCI journal list.

- The publication was identified as a **book** (b) (60 references). Identifying a book is very difficult and hazardous since the term ‘book’ is never mentioned. References usually indicated ‘articles published in’ and sometimes helpful ISBN numbers were indicated; but most often books and proceedings remained difficult to distinguish. Example:
  - Concept for intelligent rear light aiming for compensation of environmental effects in: Advanced Microsystems for Automotive Applications Springer 1999
  - ‘WALSH J. M. HILL R. G. & HATTON P. V. 2000. Evaluation of castable apatite-mullite glass-ceramics In “Materials for Medical Engineering” Eds. H. Stallforth and P. Revell ISBN 3 527 30123 2’
  - Genetic variability in tolerance to cadmium and heavy metals in pea (Pisum sativumL.) 2003 Euphytica Kluwer Academic Publishers ?

- Some publications were **PhD theses** (30 references)
- Some publications were **out of the scope** of the study: project deliverables, student reports…
- Several publication were doubled or even more **redundant** (28 references)
- 2 references concerned **patents** (p).

Exhibit 28 displays the proportion of suitable and unsuitable references for each research programme.

Exhibit 29 then provides a detailed view of the various unsuitable references for each research programme.

Exhibit 28 Number of selected and rejected references before processing indicators, by RTD programme

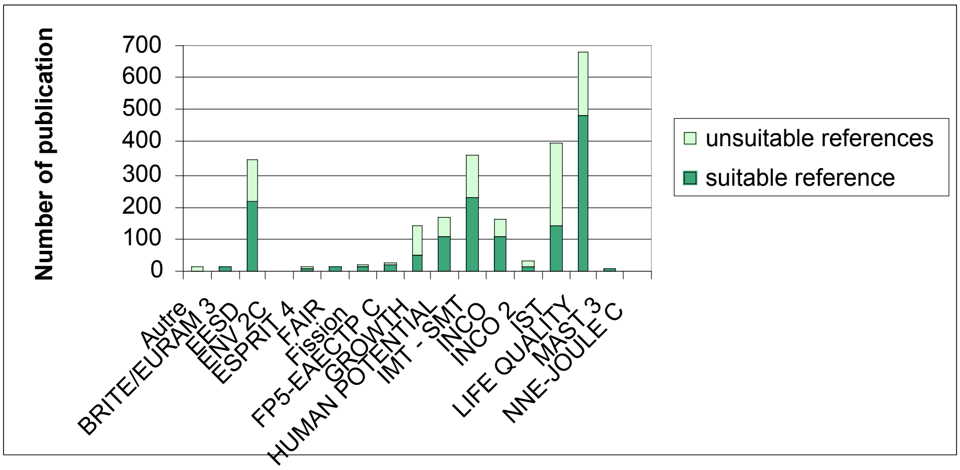
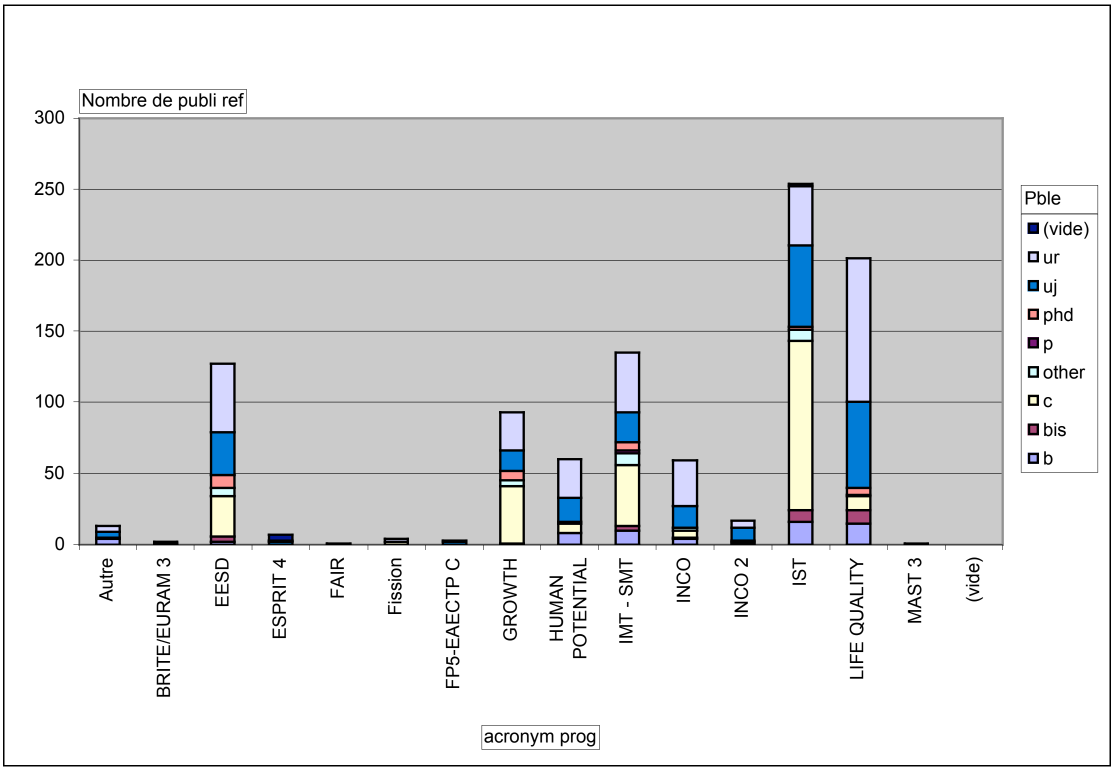


Exhibit 29 Different types of unsuitable references, by research programme



Note: ur = uncomplete reference; uj = unknown journal (absent from SCI database); phd = thesis document; p = patent; c = conference; bis = redundancies; b = book; other = other publication types of publications (reports…)