

ZEW Conference  
**Evaluation of Government Funded R&D Activities**  
Vienna, 15<sup>th</sup> and 16<sup>th</sup> May 2003

**"The Evaluation of Public Research Institutions in Italy: Comparing Different Approaches"**

Emanuela Reale

CNR CERIS (ex CNR ISPRI)  
Via dei Taurini, 19  
00185 Rome Italy  
Phone: +390649937853  
Fax: +39064463836  
e.mail: reale@isrds.rm.cnr.it

**Abstract**

One of the most innovative policy measure drawn up in Italy in recent years is linked to the introduction of *ex post* performance evaluation for both universities and public research agencies, and of *ex ante* evaluation procedures for the selection of projects and programmes.

The aim of the paper is to discuss the characteristics of the different methodologies for the evaluation of research experimented in Italy by the two bodies in charge of the assessment of public funded research institutions, and their effects on public policies. The two bodies are the National Committee for the Evaluation of the University System (CNSVU) and the National Committee for the Research Evaluation (CIVR) for the other public research organisations. Furthermore the ongoing implementation of the national evaluation system is discussed, as described in the forthcoming Government Guidelines for the Research Evaluation, designing an unified framework for all the Government funded R&D activities.

**1. Introduction**

The great emphasis given to the evaluation of research in most industrialised countries starting from the eighties, derives from various causes. First of all, Government become more and more concerned toward the optimisation of the resource allocation processes, due to the budget constrains as well as to the need of accountability before the taxpayers (Oecd, 1997, Id. 1998). This concern has been translated into a broadly application of the principle of "value for money". Secondly, public research organisations (mainly universities and university-related agencies) have been pushed to use evaluation processes for promoting a better understanding of their performance and for enhancing their accountability.

Evaluation commonly is regarded as a mean for interpreting or judging the quality, efficiency, relevance, viability, and effectiveness of university and university-related research. Evaluation helps to optimise research institutions, and emphasise the application of explicit, rational criteria for decision-making and policy-making. (Campbell D.F.J. 2000) In this paper, the terms evaluation and assessment are taken as equivalent even if the meaning is not properly the same (Hills P.V. and Dale A.J., 1995).

During the last decade we can face a rapid growth in the evaluation exercises concerning public funded research. Evaluation has been put on the agenda of most public organisations as one of the central process to enhance their performance. Systemic evaluations, based on a formal procedure that use the combination of input and output to determine the funding allocation, and applied to

non-research institutions, have progressively been extended into the research structures, where evaluation was traditionally linked to peer review processes.(Hansson, 2000).<sup>1</sup>

The aim of the paper is to present the characteristics of the different methods for the evaluation of research experimented in Italy by the two bodies in charge of the assessment of public funded research institutions, and their effects on public policies. The two bodies are the National Committee for the Evaluation of the University System (CNSVU) and the Committee for the Evaluation of Research (CIVR) for the other public research organisations.

Moreover the ongoing implementation of the national evaluation system is discussed, as described in the forthcoming Guidelines for the Research Evaluation, designing a unified framework for all the Government funded R&D activities.

## 2. Evaluation methods

Evaluation typically refers to analysis carried out on the basis of quantitative and qualitative data and information and/or peers judgements, aimed to understand the nature and the extend of the output of a research unit or institution, in relation to the available resources. Ex-ante evaluation assess the potential importance of the research and the possibilities to be successful, while ex-post evaluation focus on the assessment of the output produced and its impact on economy and society.

For the evaluation of research, peer review seems an unavoidable method, since it presents a minor level of drawbacks than interpreting data and indicators (even if it has higher costs). Peer review strengths ground on complexity (the large set of information that can be taken into account, bigger than those provided by indicators). Experts can conduct analysis more complexes than these allowed by indicators, while its weakness ground on subjectivity (the composition of the panel can bias the peer judgement. Furthermore, in some specific field, it is difficult to find real independent peers, Kostoff R.N., 1997, Van den Beemt, 1997). Peer review still represents the standard approach to the evaluation of research, and decisions about allocating resources for science. Experts reviewing the work of their colleagues should rightly be the basis of the evaluation of research.

Many factors challenged the peer review system, as well as all the evaluation procedures. One of these factors can be identified in the growth of the amount of public resources devoted to the research effort and the need of a more objective measures for research assessment. The need for a different consideration, within the evaluation exercises, of the interaction between research and the society, comes from the critique to the Mertonian idea of a system of science, in which the social, personal, organisational and political factors do not play any role (Merton R.K., 1968, Merton and Zuckerman, 1971, Whitley R., 2000, Latour, 1987, Fuller, 2000), and the overcoming changes in the mode of knowledge production (Gibbons and alii, 1994, Etzkowitz and Leydesdorff, 2000).

All the quoted elements reinforced the experimentation of different methods (programme evaluation, case studies, benchmarking), and stressed the importance of quantitative analysis (publication and citation analysis) as a supplement to peer review. Indicators give a judgement based on measures, which means quantitative data or information. The strengths of this method ground on its objectivity and its minor costs, while its weakness ground on the superficiality of the judgement (van Raan R.T.H., 1988, Rinia E.J. et alii, 2001) and on the difficult to capture the real value of the research activities.

Furthermore, a “modified peer review” model has been developed in the Scandinavian countries, “where evaluation criteria, evaluators and specific procedures are open for negotiation”, and the evaluation of research results from a compromise between the universities and the Ministry of Education (Hansson, 2000, Foss, Borum, 1999).

Thus, it is widely accepted that a completely successful model for evaluating the research activities does not yet exist. Actually, we can face different practices and experiences within the countries, even in those countries belonging to geographical areas, such as the European Union, where a great

---

<sup>1</sup> There is a difference between research evaluation and evaluation of research grounded on the fact that the former is an integrated or internal part of research, while the second is an external procedure based on various techniques used to assess individuals, institutions and research proposals (Hansson, 2000).

movement toward a further integration of research efforts can be identified. Moreover, even within the same country we see differences in the adopted models for evaluation of research, depending on the typology of institution involved, its goals, functions and tasks.

### 3. Evaluation of research in Italy: the Universities

The overall organisation of the Universities and the mechanism for resources allocation has been challenges, in the nineties, by the need of introducing a new performance-based ex-post evaluation system, linking the amount of funds to the results coming from evaluation of both teaching and research activities.

The general characteristics of the University system in Italy are:

- the largest part of universities are public, and a ratio out of 80% of the budget come from Government;
- all the Universities carry out both teaching and research activities, without the possibility to avoid one function for the other;
- the Ministry of Education, university and research (Miur) provides the core funding, by allocating it according to the number of teaching and research staff. This core funding represents the main ratio of the universities' public funding for research;
- other public funding come from grants distributed by the Miur on a competitive basis, while funds coming from public research agencies as CNR are now playing a minor role. The grants are allocated for specific national research projects, which are accessible also to non-university research institutions.

The pressure for a greater accountability of the universities started at the beginning of nineties, after the approval of the new law on the autonomy of universities (Biggier and Scarpitti, 1998, Boffo and Moscati, 1998). The law established a national body for the evaluation of the universities (*Osservatorio permanente per la valutazione del sistema universitario italiano*) and the constitution of "Units for the internal evaluation" within each University. The actions carried out by the *Osservatorio* were mainly devoted for the building up of a set of data and indicators about the university activities, also defining the Units commitment for supporting the *Osservatorio* tasks, by supplying data and analysis as requested.

As to the evaluation of research, a central role has been played by the CRUI, the Italian Standing Conference of Rectors, by stimulating "reflection and dialogue on issues related to the establishment with the universities of periodical evaluation practices. CRUI provided assistance to the universities in setting up an internal evaluation system, as well as proposing and testing possible procedure and operations. ... Thus, CRUI played a leading role in defining evaluation procedures and methods and in diffusing the culture of evaluation among universities" (Boffo and Moscati, 1998).

In 1999 the *Osservatorio* was transformed into a new body, the CNVSU, as part of the Miur. The CNVSU has a set of complex tasks, and it is aimed to carry out the overall evaluation of the universities. The law obliged the Units for the internal evaluation for providing all the data and information yearly requested by the CNVSU. Although the context for the universities' evaluation is changed, the CRUI maintains its leadership role for the evaluation of research, while the CNVSU activities were mainly focused on the evaluation of education functioning and teaching performance (both degree courses, and post-degree courses). The effectiveness of the CNVSU effort for the evaluation of the universities' educational performance is guaranteed through the reception of the evaluation results into the Ministerial decision-making.

By the end of nineties, the CRUI proposed a complex method for the evaluation of research carried out within universities, aimed to assess the disciplinary macro-sectors (CRUI, 1999).

The basic principles of the proposed method were:

- the level of evaluation is the university department or institute. The research activities to be evaluated are those related to the prevalent disciplinary macro-sector of the department/institute under evaluation;
- the assessment exercise is referred only to the efficiency and effectiveness of the macro-sector, and it does not imply any evaluation of individuals;

- the evaluation process is divided in two phases. The first is the self-evaluation phase, developed by the Units for the internal evaluation, which aim is to produce a Report on the scientific performance of the disciplinary macro-sector. The second phase is the external evaluation, carried out by a Committee of peers. The Committee would review the Reports, controlling both the data and the evaluation process results, through *ad hoc* auditing and visiting. The Committee should express a critical judgement on the Report, supplying suggestions and recommendations for bettering the scientific performance.

The proposed method use both quantitative and qualitative approaches for obtaining more in-depth of the research effort. The self-evaluation process focuses the analysis on the research objectives and on the activities aimed to gain the established goals. The large amount of indicators (70) identified by the CRUI refer to the input resources, the process, the output of products and the context in which the research activities have been developed. The external evaluation focuses on criticism of the self-evaluation process, by controlling both strengths and weakness within the macro-sector. The process was intended as an experimental phase to be implemented and improved through the simplification of indicators and the definition of a method for comparing the different disciplinary macro-sectors.

After the CRUI proposal, a certain number of universities developed evaluation practices, but the chosen methods were not exactly the same as the CRUI proposal. As a consequence, the results of the experiences cannot be compared and no conclusions can be drawn both on the effectiveness of the proposed model and on the effects of the evaluation results. Anyway, looking at the experiences of a few universities (namely the University of Florence, Bologna, Naples Federico II, Padova, Pavia, Siena, Udine, Catania)<sup>2</sup> we can face some regularities in the adopted practices, that are:

- the Reports elaborated by the Units for the internal evaluation always use different indicators for the analysis of scientific disciplines and the humanities and social sciences. Bibliometric analysis (both impact factor and citation analysis) are limited to a restrict number of disciplinary sectors;
- the Reports in many cases look more like a monitoring practice than evaluation exercise, since the effort is mainly devoted to the collection of data and indicators, without developing performance or impact analysis. Only in some cases, indexes for representing the value of research activity have been elaborated, for instance, for assessing the departments ability to perform research, scientific development and dissemination of results (Carotenuto et alii, 2001);
- the differences between universities for carrying out a good evaluation exercise are related to the existing level of confidence with the evaluation processes. The level of confidence is mainly linked to the amount of funding coming from external sources (both national and international funds, specially European Union funds), and to the magnitude of the scientific macro-sectors, more used to undergo evaluation procedures.<sup>3</sup>

Thus, the comprehension of the effects on public policies coming from the evaluation of research within the universities is not yet a simple affair. We can easily statue that evaluation produces a substantial improvement of the universities acceptance of the ex-post performance assessment. In some cases we also face direct effects on the internal resources allocation (Carotenuto *et alii*, 2001). However, it is important to underline that the acceptance of the CRUI method by the universities, as well as the experimentation of self-determined research evaluation exercises, occurred on a voluntary basis, and there is not a formal linkage between evaluation of research and resources allocation.

The CRUI launched recently a new initiative for analysing the universities' scientific production, by using bibliometric indicators (Breno *et alii*, 2002). Three indexes have been elaborated for

<sup>2</sup> The Reports of the Units for the internal evaluation are available at the respective web pages, that are: [www.unifi.it](http://www.unifi.it), [www.unibo.it](http://www.unibo.it), [www.unipd.it](http://www.unipd.it), [www.unipv.it](http://www.unipv.it), [www.unisi.it](http://www.unisi.it), [www.unict.it](http://www.unict.it).

<sup>3</sup> The literature also underline the importance of the Rectors' attitude and personality in favour of research assessment (Boffo and Moscati, 1998).

comparing scientific production and productivity among the universities: the citation impact (CI), the presence index (PI) and the productivity index (PII).

The characteristics of the CRUI analysis are:

- source of data for publications: ISI-SCI (expanded). Publications include articles, notes, reviews, proceedings, following the ISI classification of products;
- descriptors:  $D_1$ =number of human resources (professors and teachers) of the Italian universities by the September 2001;  $D_2$ =number of ISI publications of the Italian universities in the period 1995-1999;  $D_3$ =number of citation obtained in the same period by the publications of the ISI National Database;
- indicators:  $(PII)=D_2/D_1$ ;  $(PI)=D_3/D_1$ ;  $(CI)=D_3/D_2$ .
- anomalies have been removed in the sample for: *a*) universities with a low number of human resources ( $D_1 <_{\min}(0,03 * D_{1,\max}, 20)$ , where  $D_{1,\max}$  is the maximum value of  $D_1$  in the scientific sector; *b*) universities with a low number of publications ( $D_2 <_{\min}(0,03 * D_{2,\max}, 20)$ , where  $D_{2,\max}$  is the maximum value of  $D_2$  in the scientific sector. In both cases the hypothesis is that the universities with no significant levels of  $D_1$  and  $D_2$  in some sectors, do not have a relevant research effort in that sectors. Thus, the research effort is relevant if in the selected area there is a number of professors and researchers higher than 3% of  $D_{1,\max}$ , but no less of 20. The same rule for relevance is applied to the number of publications, which should be higher than 3% of  $D_{2,\max}$ , but no less of 20.

The comparison between universities separates the universities with Medicine courses and the universities without Medicine courses, for differences in publishing trends. Table 1 and 2 show the results of data analysis. The differences depends mainly on two factors: the robustness of hard sciences within the university (hard sciences are best represented by the ISI data than social sciences and humanities) and the localisation of the structures, with a worse position of the universities located in the South of Italy. The size of universities, measured in terms of human resources (professors and researchers) is not an element influencing the best performance.

The analysis had a great impact on the public debate about the Italian research system, and demonstrates the permanence of a high level of attention about the accountability of the national university system from the CRUI.

#### **4. Evaluation of research in Italy: the university-related agencies**

In 1998, the law n. 204 reformed the Italian research central organisation and created the CIVR, for the assessment of the non-university public funded research organisations.<sup>4</sup> The law established also the settlement of one Internal Evaluation Committee-CIV within each public research agencies under the CIVR control.

CIVs are panels of experts, nominated by the agencies themselves on the basis of criteria established by the CIVR. The CIVs aim is to develop systemic evaluation of the whole research organisation performance. The panels are composed from 5 to 7 members; they include both experts in the specific discipline or sector of activity (peers in the strict sense), as well as experts in the economic assessment of the internal management. In some cases also potential users of the research activities have been nominated. The committees should include a certain number of components coming from abroad; the impartiality of the judgement is guaranteed by the absence of institutional relations of the CIVs members with the agency. The CIVs work should fit with the requests coming

---

<sup>4</sup> The Committee institutional tasks are aimed towards:

- the diffusion of the evaluation culture within the country,
- the setting up of general criteria and indicators for the ex post assessment of the public funded research activities (those carried out by both public and private structures),
- the definition of the conditions to be applied for the composition of the Internal Evaluation Committees - CIVs of the public research agencies.

from the CIVR, complying with the set of evaluation criteria that have been requested. Furthermore, CIVs can identify and explore other relevant criteria for the agencies' assessment.

CIVR developed a three-year work (1999-2001) carrying out ex post evaluation exercises for the most important Italian non-university research agencies (eight major agencies), and monitoring the state of the art of research evaluation in Italy.<sup>5</sup> The group of research agencies examined, mainly act as performer of research activities rather than as funding agencies. In some cases (see CNR, for instance) the agencies have also a funding function, but it is less relevant than the research function. The limits of the CIVR results are due to the fact that data and figures cover only three years, namely 1999, 2000 and 2001. So, it is not possible to measure changes in time of the agencies' behaviours linked to the evaluation process. Modifications of the actors' behaviours should be identified, in these cases, through the peers' analysis.

The evaluation exercises developed by the CIVR covers the main public research agencies outlined in Table 3. The mentioned agencies represent a ratio out of about 74% of the total Italian investment in public research agencies (1.631,4 million Euro in 2000, on a country total of 2.208 Euro for the research agencies), as well as a percentage out of 58,6% of the total public national expenditures (2.784 million Euro, universities not included, CIVR, 2002).

From a methodological perspective, CIVR analysis try to combine both external reviewing and quantitative analysis through a set of selected indicators (European Commission, 1997). CIVR asks the agencies for carrying out two different exercises.

Firstly, a self-evaluation exercise, make by the agencies themselves, that are committed for a critical review of their performance according with a set of proposed criteria.<sup>6</sup> All the criteria have been operationalized by a group of indicators suited to measuring the agency performance. Criteria and indicators have been discussed and agreed with the interested institutions and with the CIVs; the use of the same criteria should assure the viability of the exercise and a certain level of comparability of the common trends between different types of institutions.

Secondly, a CIV evaluation exercise, based on both the self-evaluation results and other knowledge activities autonomously decided by the Committee (local visits, auditing, special meeting, other expert views, indicators, etc.).

Then, CIVR would revised both the exercises,<sup>7</sup> comparing the coherence of the obtained results for:

---

<sup>5</sup> In our discourse we refer to a group of public research agencies, that are entities traditionally performing university-related research: "referring to standardised Oecd terminology, university research coincides with R&D that is performed by the higher education sector; and university-related research coincides with R&D being performed by the government and private non-profit sector. ... Academic research is a science-based activity, where a major emphasis is placed on basic research and on the combination of basic and applied research." (Campbell D.F.J. 2000).

<sup>6</sup> CIVR identified a set of general criteria for the assessment exercises to be applied for each evaluation exercise (both the self-evaluation exercise and the CIV evaluation exercise). The criteria refer to:

- monitoring of the whole input, output and outcome of the institutions;
- scientific quality of the publications, measured, if possible, by the impact factor and the citation analysis;
- internationalisation of the research activities and relevance for the scientific sector;
- innovative perspective on the research programmes;
- level and characteristics of the collaborations, interactions and networks;
- impact on the socio-economic environment;
- focalisation of the selected objectives with the mission of the evaluated institute;
- capacity to attract external financial and human resources;
- new management capabilities for sustaining and encouraging the research effort.

<sup>7</sup> The CIVR grounds the judgements about the CIV Reports on three main criteria, namely the clearness of the results achieved (in terms of the capability to identify points of excellence as well as problematic issues, measuring them by comparing the agency performance with the performance of other similar agencies abroad, and formulating recommendations to overcome the problems and to maintain the scientific leadership in the areas of excellence), the transparency of the evaluation processes, and the reliability of the adopted methodology. On the other hand, the CIVR judgements on the self-evaluation exercises take into account the capability of the agencies for:

- supplying figures, indicators and information as requested,
- showing strengths and weakness within the internal organisation as well as in the research planning,
- identifying innovative perspectives within the planned research activities,
- using the evaluation process for the internal decision making.

- a) the collaboration and networking of the research agencies with other scientific actors;
- b) the trend towards the publication on international journals (both as absolute values and as equivalent values);
- c) the visibility and leadership at the international level,
- d) the capacity for attracting external resources,
- e) the interaction with the socio-economic environment,
- f) the propensity towards patenting and spin off.

CIVR highlights explicit difficulties for some agencies in developing in depth evaluation exercises, as well as for complying with the recommendations coming from the CIV assessment. Problems arise also by using qualitative indexes (impact factor and citation index) for measuring the scientific output. So, neither performance comparison between the agencies has been made, neither ranking list on the basis of the selected criteria. Thus, the final judgement on the agencies performance was based on the identification of points of strengths and weakness in pursuing their institutional mission, mainly through the qualitative judgements provided by the CIVs.

Anyway, we faced a great movement for building up the evaluation exercise within the agencies, due to the fact that they are obliged to undergo the CIVR examination. The movement involved both agencies where an internal evaluation process was present (such as INFN and INFN) and agencies without any previous experience in systemic performance assessment.<sup>8</sup>

Moreover, the CIVR Report impact on public policies for research and development in different ways. The Miur made reference to the CIVR results as justification for the ongoing reform of the public research system. At the end of 2002, CIVR was committed for building up the Miur guidelines (criteria and process) for the evaluation of all the research funded by the Ministry itself: universities, research agencies, and national research projects.

## 5. The new Guidelines for the evaluation of research

The Guidelines was settled by merging the different national experiences on evaluation carried out within universities and the CIVR activities. The proposed system includes the overall assessment of the Miur funded R&D: *ex post* performance evaluation of structures (both universities and research agencies under the Miur control), *ex ante* and *ex post* evaluation of national research programmes.

Furthermore, the document is conceived as a proposal of rules and procedures for the evaluation of all the Government funded R&D, as well as a base for public discussion on research evaluation issues.

Focusing the attention on the process organisation, we can face the following characteristics. The exercise occurs every three years, and it take three years to be completed. The estimated total costs are of 8.545.000 Euro.

The basic organisms are the Units for internal evaluation of the universities, and the CIVs of the research agencies. They are nominated autonomously by the research structures. Both the Units and the CIVs should elaborate a three-year Report containing the evaluation of the structure (the so-called self-evaluation) on the basis of a set of criteria indicated by the CIVR. They should also provide validate data and indicators about the structure, as requested by the CIVR.

Every structure under evaluation should select a number of products, equivalent to the 50% of the average number of full time equivalent researchers belonging to the structure, during the three years of the evaluation exercise.<sup>9</sup> Given the actual size of the Italian public research system in terms of researchers, CIVR estimated to evaluate 25.000 products.

Panels are the organisms in charge for the evaluations of products and projects. There is one Panel for each disciplinary area,<sup>10</sup> one Panel for each national project and one panel for each special area. Special areas should be identified by the CIVR at the beginning of the evaluation exercise, also taken into account the priorities of the National Research Plan and of the European Research Programmes. The number of special areas should not exceed the number of disciplinary areas. For areas or projects characterised by a specific heterogeneity and/or by a great number of products to be assessed, the CIVR may constitute, within the existing Panels, sub-panels with specific competence. Panels are composed by a minimum of 5 to a maximum of 9 peers. The selection of peers to be included in the Panels is made by the CIVR, which is integrated, for this purpose, with external observers, designed by the scientific community. The

---

<sup>8</sup> The movement implies the growth of the agencies requests for CIVR auditing, the organisation of national conferences, seminars and forum for discussing issues related to the evaluation of research, the constitution of task forces within the agencies for supporting the CIVs activity. In some cases, as CNR, there was an adaptation of the agencies priority setting to the CIVR indications (CNR cuts-off of the funding of *extra muros* research activities for the low level of the available resources, it decentralises the administrative personnel, and sets up measures for reinforcing patenting and spin-off).

<sup>9</sup> Researchers include stabilised professors and researchers, as well as fixed term contract researchers. It does not include fellowships, doctoral students and other training personnel. Products that can be selected are: the articles, the books and chapters of books, the patents, the projects, the compositions, the designs, the performance, the exhibitions, the artefacts and the artistic works. Bibliometric indexes (impact factor and citation rate) should be included in the description of the product, if applicable.

<sup>10</sup> The disciplinary areas identified are 14 and correspond with the Universities disciplinary macro-sectors

observers should control the transparency of the selection process, during all the phases in which it is articulated, as suggested by the European Union Guidelines for Evaluation Procedures.<sup>11</sup>

Panels should use experts, even coming from abroad, for reviewing the selected products. Almost two experts should revise each product. It means that 55.000 experts' reviews should be forecast. Experts judge the quality, the relevance, the originality and the international competitiveness of the product, providing the rating on the basis of a four-step scale: excellent, good, acceptable, poor. In case of experts' disagreement, the Panel should provide the definitive judgement.

After collecting the experts' advice, the Panels draw up the Final Report for each Area or Project, and a ranking list of the products.<sup>12</sup>

CIVR is the main organism in charge of the evaluation, with a complex set of tasks, including:

- the direction and co-ordination of all the process,
- the identification of criteria for the CIVs and the Panel composition,
- the evaluation of the Reports coming from the CIVs, from the Units for internal evaluation of the universities, and from Panels,
- the writing of the Final Evaluation Report for each structure and project,
- the proposal of solution for linking the evaluation results and the resource allocation.

The CIVR process for evaluating the research structure is based on the analysis of the CIVs and of the Units Reports, as well as on the results coming from the Panels' examinations. Further analysis, especially those devoted to the assessment of the socio-economic impact of the research results, could be carried out by using the set of descriptors collected for each structure or projects, as well as by using the selected indicators. Pursuing its knowledge objective, the CIVR may adopt the quantitative technique best suited for its purpose (case studies, benchmarking, bibliometric analysis, cost-benefit analysis, etc.).

The overall organisation of the process is described in Fig. 1, while Table 4 provides a comparison between the proposed system and the Research Assessment Exercise RAE, used in United Kingdom for evaluating the research performance of the university system.

The differences with the previous experimented methods are mainly linked to the circumstance that not all the products should be evaluated, but only a selection of products made by the structures themselves, probably the best ones.

Secondly, the level of evaluation is not the departments or the institutes, but the entire research structure, university or research agency. Thirdly the exercise should evaluate the disciplinary macro-sectors, the special areas and the national projects, allowing the identification of sectors and structures of excellence. Finally, the maintenance of both peer reviewing and quantitative analysis, should assure a better inside of the value of the research, improving the public system accountability also for research results more linked to the societal needs.

The proposed exercise seems to avoid some disadvantages of the performance-based evaluation systems, namely the high cost and the tendency towards "publish or perish", and the "publication inflation", with researchers seeking to split the results of their work into the least publishable units" (Geuna and Martin, 2001).

The expectation, in fact, is toward a reduction of the number of products, trying to better the scientific quality, simply by publishing on the more relevant journals or editors.

Anyway, some drawbacks still persist. First of all, the push towards publishing on the better recognised journals could create a sort of homogenisation of the research effort within the public research structures, reducing the incentives for a differentiation of the institutions' profiles. Furthermore, the proposed system could restrict the spaces for creativity and new ideas, as well as for high-risk projects, which have a greater possibility of failure than the traditional ones. Finally, since the system is aimed to provide the evaluation of excellence, it could create some misunderstanding on the effective value of the overall university research production.

---

<sup>11</sup> After the identification of the components' profiles, there is a general call for applicants. The components should be chosen after a comparative examination of the curricula, See *Guidelines for proposal evaluation procedure*, FP6, UE, 2003.

<sup>12</sup> The Panels Final Report is articulated into three parts. The first part is the Consensus Report, which re-examines the experts' judgements and the bibliometric indexes, if applicable. The second part is the ranking list of the products per areas, special areas and projects. The third part is the final one, providing an analysis of strength and weakness of the areas and project, and suggesting measures for bettering the general performance. A specific section of the Report should be dedicated to the patent analysis.



## 6. Conclusions

Italy has remained for a long time a country where knowledge production and evaluation were not linked processes, and the research assessment was not used by policy makers for resource allocation decisions.(Oecd, 1992).

The first action for the evaluation of research was related to the establishment of the evaluation process for the universities. It started in 1993, and substantially implemented in 1996, by the setting up of the National Committee for the evaluation of the Universities, CNVSU, which aim is to evaluate both teaching and research activities of the universities.

After the subsequent reform of the 1998-1999 (De Marchi *et alii*, 1998), evaluation is more and more becoming a central issue for the decision-making. As the autonomy of research organisations was reinforced, evaluation was enhanced, and a National Committee for the Research Evaluation, CIVR was created, for the non-university public funded research. In the Government intention, the activities of the two committees, for what concern the assessment of the research effort, should be integrated through a combination of the adopted approaches.

Looking at the Italian experience, the factors reinforcing the stabilisation of a new attitude towards the evaluation of research have been the presence of an obligation to the evaluation exercise, and the strong Government commitment toward improving the accountability of the Government funded R&D.

By the contrary, the absence of a formal linkage between evaluation and resources allocation has been a factor impeding the effectiveness of evaluation results, reducing the possibility of feedback effects, almost within universities.

The new Guidelines merge the different experiences of universities and public agencies evaluation of research, also including some aspects of other evaluation systems existing abroad.

The process include both quantitative methods for evaluating the institutions' performance, and the peer reviewing for judging the quality, relevance and originality of the selected products. The described characteristics should avoid some disadvantages of the research performance evaluation. More specifically, the high cost, the "publish or perish" tendency, the absence of attention to research themes linked with social needs. On the other hand, the CNVSU effort for evaluating the universities educational performance, guarantee the maintenance of a balance in the internal priority setting for both teaching and research activities. In this perspective, it is important that results coming from CNVSU and CIVR evaluation exercises are both considered for keeping a real knowledge of the universities' role within the national research system.

Anyway, the effectiveness of the Guidelines process should be assured by the parallel introduction of a performance-based funding system, grounded on the evaluation of research. As to the universities, the shift should probably be from the actual system, based only on educational size, to an "hybrid" system, that is one based both on performance and on education size (Geuna A., Martin Ben R., 2001). As to research agencies, the future is a more open question: would they pass completely to a performance-based resource-allocation system, or a sort of mixed system could be adopted? A hybrid system is more suitable, in our view, even for university-related agencies. It would overcome the cited disadvantages of a performance-based evaluation of research, due to its capacity to maintain larger spaces for the individuals creativity and non traditional research approaches, as well as for limiting the Government interference into the setting up of the agencies research agenda.

## References

Biggieri L., Scarpitti L.(1998), *Evaluation in the Italian University System*, Paper presented at the International Conference on: "Evaluation: Profession, Business or Politics?", Rome, October 29-31

Boffo S., Moscati R.(1998), Evaluation in the Higher Education System: many tribes, many territories ... many godfathers, *European Journal of Education*, Vol. 33, No. 3, pp. 349-360

Breno E., Fava G.A., Guardabasso V., Stefanelli M. (2002), *La ricerca scientifica nelle università italiane. Una prima analisi delle citazioni della Banca Dati ISI*, CRUI, Roma

Campbell D.F.J. (2001), *Quality improvement and the evolution of academic research systems: five hypothesis on the evaluation of university research*, Proceeding from the 2000 US-European Workshop on Learning from Science and Technology Policy Evaluation, Shapira P. and Kuhlmann S. (eds.), at [www.cherry.gatech.edu/e-value](http://www.cherry.gatech.edu/e-value) or [www.isi.fhg.de/ti/departm.htm](http://www.isi.fhg.de/ti/departm.htm)

Carotenuto G., Lapegna M., Zollo G., Di Donato A., Nicolais L.(2001), Evaluating research performance: the strategy of the University of Naples Federico II (Italy), *Higher Education Policy*, 14, pp. 75-90

CIVR (2002), *Relazione annuale 2000-2001*, Roma, Miur-Ministero Istruzione, Università e Ricerca. The English version of the Report is available at [www.civr.it](http://www.civr.it)

CRUI (1999), *Metodo di valutazione della ricerca svolta presso strutture scientifiche universitarie nell'ambito del macro-settore scientifico-disciplinare prevalente*, Roma

De Marchi M., Poti B. M., Reale E., Rocchi M., Scarda A.M.,(1998) *Changing Structure, Organisation and Nature of PSR Systems Italy, Final Report*, (UE Project European Comparison of Public Research System), Rome, June

Etzkowitz H., Leydesdorff (2000), "The dynamics of innovation: from national systems and "Mode 2" to a Triple Helix of university-industry-government relations", *Research Policy*, Vol. 29, p. 109-123

European Commission (1997), *Second European Report on Science and Technology Indicators*, Brussels

Foss H.H., Borum F. (1999), The construction and standardisation of evaluation. The case of Danish University Sector", *Evaluation*, Vol. 5, n. 31, pp. 303-329

Geuna A., Martin Ben R. (2001), *University research evaluation and funding: an international comparison*, SPRU Paper No. 71, forthcoming on *Minerva*

Gibbons M., Limoges C., Nowotny H., Schwartzman S., Scott P., Trow M., (1994), *The new production of knowledge. The dynamics of science and research in contemporary society*, SAGE Publications.

Hansson F. (2000), *The peer review system in the changing knowledge production system*, presented at Evaluation 2000, AEA-Conference, November 1-5, Waikiki, Hawaii

- Hills P.V. and Dale A.J. (1995), "Research and Technology Evaluation in the United Kingdom", *Research Evaluation*, Vol. 5, 1, pp. 35-44
- Kostoff R.N. (1997), "Peer review: the appropriate GPRA Metric for research", *Science*, Vol. 277, p.
- Oecd (1987), *Evaluation of research. A selection of current practices*, Paris
- Ocde (1991) *Examens des politiques nationales de la science et de la technologie. Italie*, Paris
- Oecd (1997), *The Evaluation of Scientific Research: Selected Experiences*, Paris, Oecd
- Oecd (1998), *Policy evaluation in innovation and technology. Towards best practices*, Oecd, Paris
- Rinia E.J., van Leeuwen Th.N., van Vuren H.G., van Raan A.F.J (2001), "Influence of interdisciplinarity on peer review and bibliometric evaluations in physics research", *Research Policy*, 30, p. 357-361
- Van den Beemt F.C.H.D. (1997), *The right mix by peers as well as by highly qualified persons (non peers)*, Australian Research Council Commissioned Report on "Peer review process", n. 54, p. 153-164
- Van Raan R.T.H.(1988), *Handbook of Quantitative Studies of Science and Technology*, Amsterdam, North Holland

<i>Table 1 - Impact index, productivity index and presence index of the Italian universities(a).</i>				
<i>Years 1995 - 1999.</i>				
Universities with Medicine courses.				
<i>University</i>	<i>Impact (CI)</i>	<i>Productivity (PII)</i>	<i>Presence (PI)</i>	
Ancona	3,3	4,1	13,5	
Bari	3,9	2,9	11,1	
Bologna (b)	4,1	4,1	16,9	
Brescia	6,6	5,0	33,1	
Cagliari	4,2	2,8	11,8	
Catania	3,6	2,6	9,3	
Chieti (b)	4,5	3,1	14,1	
Ferrara (b)	4,5	4,6	20,4	
Firenze	4,6	4,1	18,6	
Genova (b)	4,5	3,9	17,3	
L'Aquila	3,7	4,3	15,8	
Messina	3,0	1,8	5,3	
Milano	4,9	7,4	35,9	
Milano Cattolica (b)	4,2	2,9	12,1	
Modena Reggio Emilia	4,0	4,5	17,6	
Napoli I Federico II	3,8	3,0	11,4	
Napoli II	3,7	1,8	6,8	
Padova (b)	5,0	5,3	26,7	
Palermo	3,3	1,5	5,1	
Palma	4,3	3,4	14,6	
Pavia (b)	5,0	5,9	29,1	
Perugia	4,8	3,7	17,6	
Pisa (b)	4,0	4,6	18,1	
Roma I La Sapienza	4,0	3,5	13,9	
Roma II Tor Vergata	4,6	4,6	20,8	
Sassari	4,2	2,5	10,4	
Siena (b)	3,9	4,3	16,8	
Torino	5,4	4,3	22,9	
Trieste	4,4	4,4	19,4	
Udine	3,7	4,1	15,1	
Verona	5,5	5,6	31,0	
<i>Average value</i>	<i>4,1</i>	<i>3,1</i>	<i>12,9</i>	
(a) Universities not included for the limited number (less than 550) of publication in the considered period: Catanzaro, Foggia, Insubria, Milano Bicocca, Milano S. Raffaele, Piemonte Orientale, Roma Campus biomedico.				
(b) The University has not yet completed the analysis of the ISI data. Thus, the indicator is provisional.				
ricerca riportate nella banca dati.				
Source: Crui, <i>La ricerca scientifica nelle Università italiane</i> , CRUI, Rome,2002.				

Table 2 - Impact index, productivity index and presence index of the Italian universities(a).  
Years 1995 - 1999.

Universities without Medicine courses.

<i>University</i>	<i>Impact (CI)</i>	<i>Productivity (PII)</i>	<i>Presence (PI)</i>
Basilicata	3,02	2,87	8,67
Calabria (b)	3,63	3,19	11,58
Camerino (b)	3,2	4,51	14,46
Cassino	2,1	1,37	2,88
Lecce (b)	2,97	4,12	12,26
Molise	2,82	2,54	7,16
Napoli Parthenope	4	1,68	6,72
Reggio Calabria Mediterranea	6,06	2,83	17,14
Roma III	3,56	4,23	15,03
Salerno	2,33	4,58	10,68
Trento (b)	3,28	6,23	20,42
Urbino (b)	6,73	3,59	24,17
Venezia (b)	3,12	4,39	13,67
Viterbo Tuscia	2,75	3,01	8,27
Valore medio	3,45	3,57	12,35

(a) Universities not included for the limited number (less than 65) of publication

in the considered period: Bergamo, Castellanza Cattaneo, Macerata, Milano Bocconi, Napoli Orientale, Roma IUSM, Roma Lumsa, Sannio, Teramo, Venezia Architecture.

(b) The University has not yet completed the analysis of the ISI data.

Thus, the indicator is provisional.

Source: CRUI, La ricerca scientifica nelle università italiane, CRUI, Rome, 2002.

Table 3. Research Agencies under the CIVR Evaluation (Year 2000)

	Total Budget*	Total Staff	<i>of which</i> Researchers	
1. C.N.R. _ Consiglio Nazionale delle Ricerche	769,0	7.377	3.650	
2. I.N.F.N. _ Istituto Nazionale di Fisica Nucleare	306,8	5.041	2.196	°°
3. I.N.F.M. _ Istituto Nazionale di Fisica della Materia	87,3	3.092	1.689	°
4. Istituto Papirologico "G. VITELLI"	0,0	13	10	
5. O.G.S. _ Istituto Nazionale di Oceanografia e di Geofisica Sperimentale	14,2	150	48	§
6. S.Z.N. _ Stazione Zoologica "A. DOHRN"	16,9	125	53	
7. ENEA - Ente per le Nuove Tecnologie, l'Energia e l'Ambiente	361,6	3.238	1.300	
8. I.E.N. _ Istituto Elettronico Nazionale "GALILEO FERRARIS"	14,0	143	60	

\*Million Euro

° 2.408 researchers coming from universities

°° 2.935 researchers coming from universities

°\* 1.427 from universities

°\*\* 1.581 from universities

§ Year 1999

Source: CIVR Annual Report 2000-2001

*Table 4 - Comparison between Research Assessment Exercise (RAE) and the proposed Italian Guidelines for evaluation of research*

	RAE	Italian Guidelines
Term of the exercise	5 years	3 years
Number of Panels	about 70	14+ <i>n</i> *
Number of Panelists	about 700	< 150
Experts	not essential	55.000 reviews
Selected products (total)	about 200.000	about 25.000
Evaluated products	25% of the selected	all the selected
Average rate of evaluated product per year	< 10.000	> 8.000
Selected products per researcher	out of 4	0,5**

\* Panels for Special Areas (no more than those of disciplinary areas)

\*\*For ETP researchers

**Fig. 1 - The Italian Guidelines for Evaluation of Research: the Organisation of the Process**





