ISSN : 2321-9602

Indo-American Journal of Agricultural and Veterinary Sciences




editor@iajavs.com
iajavs.editor@gmail.com

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

Indo-American Journal of Agricultural and Veterinary Sciences

The Italian experience demonstrates how national peer-review research evaluation activities in the hard sciences

Ramesh.m ,K.Srinivas ,Vinay kuma .M , C . Dhanu

Abstract

Bibliometrics has been shown to be more effective than peer review in national research evaluation activities, especially in the hard sciences. In this work, we focus on Italy and use the bibliometrics-based university performance rankings list from 2001-2003 as a comparison. We evaluate the first nationwide assessment exercise based only on peer review and contrast it to other free rankings lists compiled using factors only tangentially related to performance or publicly accessible online. The findings demonstrate that the Italian review of research institutes might have been avoided entirely in the hard sciences.

Keywords
Research evaluation; bibliometrics; VTR; ranking; productivity; universities.

1. Introduction

From the 1980s forward, when governments throughout the world realized the importance of the knowledge economy, they enacted policies and programs to strengthen their own higher education institutions. In order to allocate funds fairly and encourage beneficiaries to be more productive in their own research, a growing number of countries are engaging in research evaluation exercises at the national level. Traditionally, these exercises have relied on the peer-review system, in which experts in the field analyze research outputs presented by institutions. Recent advances in bibliometrics have made it possible to use metrics that aid reviewers in their work of assessing research items (informed peer-review). In certain circumstances, like as Australia's ERA exercise in 2010, bibliometrics have totally replaced peer-review for the hard sciences.Horrobin (1990), Moxham and Anderson (1992), MacRoberts (1996, 2002), Moed (2002), van Raan (2005), Pendlebury (2009), Abramo and D'Angelo (2011), and many others have weighed in on the merits and drawbacks of peer review and bibliometric approaches. Although the literature does not provide a clear answer as to which method is superior for evaluating specific scientific products, it does show a positive correlation between peer-review outcomes and citation indicators (Serenko et al, 2011; Aksnes and Taxt, 2004; Oppenheim and Norris, 2003; Rinia et al., 1998; Oppenheim and Norris, 1997; Van Raan, 2006) and between peer-review and bibliometric rankings, whether on an individual

Institute for System Analysis and Computer Science

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013
Indo-American Journal of Agricultural and Veterinary Sciences

When used to a comparative examination of people, research groups, or whole institutions, the severe limitations of peer review become apparent. With respect to accuracy, robustness, validity, functionality, time, and costs, Abramo and D'Angelo (2011b) compare and contrast the peer-review and bibliometrics methodologies in national research evaluations. In the end, they decide that bibliometric technique is superior than expert opinion. Other implications of the peer-review system are to blame, not the fact that citation counts are more reliable than peer assessment in determining the quality of individual outputs. Peer-reviewing every piece of research that comes out of a country's research system is obviously impossible due to time and money constraints. There are evident drawbacks to using peer review instead of the bibliometric approach because of this. For starters, it makes it impossible to track productivity, the single most important indication of a manufacturing system's effectiveness. Second, the peer-review approach is substantially compromised by the fact that rankings are sensitive to the size of the subset of output that is assessed. Third, it necessitates a subset selection of products (or researchers), which is not always an efficient process and may introduce elements of distortion that undermine the validity of the peer-review method, leading to rankings that do not accurately reflect the true quality of the subjects evaluated. Fourth, it restricts the method's use since it can't be used by each and every researcher or study team. Therefore, universities do not have access to research staff performance rankings that may influence internal selective financing. Finally, in terms of the bibliometric strategy, the high costs and lengthy timelines associated with peer-review severely restrict its potential. frequency of the death penalty.

Abramo et al. (2011) compared university performance ranking lists from the first peer-review Italian research assessment exercise (VTR) to those obtained from evaluation simulations conducted with bibliometric indicators in order to measure the effects of the aforementioned limitations to peer-review methodology. Measurement of the amplitude of changes found considerable disparities for the rank of the institutions involved.

Peer-review processes have substantial upfront expenses that shift depending on how many goods are being assessed. The 2008 RAE in the UK, for instance, which assessed four products per academic, cost 12 million pounds2. It is predicted that the indirect costs to the assessed institutions are five times higher than the direct expenses due to the opportunity costs of the time spent by administrative and research employees. One recent example is the second Italian research evaluation exercise, for which direct expenses are currently projected at between 10 and 11 million euros.

Given that bibliometric methodologies provide more exact and robust findings at considerably lesser cost, one must wonder whether it is worthwhile to pay such significant sums of public money to acquire rankings that give such limited precision. In this work, we rank Italian universities in terms of performance using free, publicly available indicators that are only tangentially related to research (such as geographic or economic type), providing further evidence that peer review is inadequate for national evaluation of research institutions in the hard sciences. We use the VTR and a bibliometric based on the complete scientific output indexed in the Web of Science (WoS) as a benchmark to evaluate the accuracy and robustness of

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013
Indo-American Journal of Agricultural and Veterinary Sciences

these ranking lists. Contrary to expectations, the comparative findings demonstrate that the VTR ranks are not more precise than the free alternatives.

The first is a ranking of schools from northernmost to southernmost, based on the value of their latitude. Two more sets of rankings are obtained from economic data, including gross domestic product by area and research spending by region and resident. We conclude by comparing two worldwide university rankings: one derived from the SCImago Institutions Ranking (SIR) World Report3 and the other created by the Italian socioeconomic research institution CENSIS4.

We then provide examples of the procedures used to arrive at the rankings that constitute the meat of the article. The analyses and their outcomes are presented in Section 3, and the report's takeaways are summarized in Section 4.

Creating the numerous lists of rankings

1.1 The VTR evaluation

The Italian Ministry of Universities and Research (MIUR) initiated its first Research Evaluation exercise, VTR, covering the years 2001–2003, in December 2003. The VTR was tasked to a national body called the Directory Committee for the Evaluation of Research (CIVR). Research and development efforts at 77 universities, 12 governmental research institutes, and 13 commercial research institutions (the latter of which participated at their own cost) were to be evaluated by the evaluation system. Only university rankings are examined in this research.

There are 370 different disciplinary sectors (SDSs) at Italian universities, organized into 14 different UDAs, 8 of which are in the hard sciences. Research performance rankings for each UDAs were supplied by CIVR. First, the CIVR assembled expert committees to evaluate each UDA. The panels then required universities to provide research results on their own initiative, with each institution submitting results equal to one quarter of its total research staff during the monitoring period. Only articles, books, and chapters of books, congress proceedings at the national and international levels, patents, and designs, as well as performances, exhibits, and works of art were considered acceptable outputs. Next, the panelists evaluated the findings and assigned final ratings of "excellent," "good," "acceptable," or "limited" to each product. The CIVR assembled panels of 183 highly-regarded peers who regularly sought advice from subject matter experts. Each institution in each UDA6 was ranked using the following quality rating R based on the following criteria: quality, relevance and innovation, international scope, and capacity to promote international competitiveness.

$$R =^{1} \cdot (E + 0.8G + 0.6A + 0.2L) \quad [1]$$

Where:

E; G;; L= numbers of “excellent, good, acceptable” and “limited” outputs submitted bythe university in the UDA

T = total number of outputs submitted by the university in the UDA

A final report ranks universities based on their results under the quality assessment rating. As an example, Table 1 shows the ranking list of the top 10 Italian universities basedon R, in the UDA “Mathematics and computer science”.

The magnitude of the VTR effort is suggested by a few facts: the evaluation included 102 research institutions and examined about 18,000 outputs, drawing on 183 panelists and 6,661 reviewers, with the work taking almost two years and with direct costs mounting to
3.5 million euro.

Complete list accessible at http://cercauniversita.cineca.it/php5/settori/index.php, last accessed on Sept. 5,2012.
[6] http://vtr2006.cineca.it/index_EN.html, last accessed on Sept. 5, 2012.

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013
Indo-American Journal of Agricultural and Veterinary Sciences

University	Selected outputs	E	G	A	L	Rating	Category Ranking (%)
Sissa	3	3	0	0	0	1.000	100
Sannio	1	1	0	0	0	1.000	100
Rome "Tor Vergata"	23	17	5	1	0	0.939	96.15
Milan	28	17	10	1	0	0.914	92.31
Bari Polytechnic	7	4	3	0	0	0.914	92.31
Milan Polytechnic	25	16	7	2	0	0.912	90.38
Insubria	6	3	3	0	0	0.900	86.54
Verona	4	2	2	0	0	0.900	86.54
Pisa	42	22	18	2	0	0.895	84.62
Turin Polytechnic	19	9	10	0	0	0.894	82.69

Table 1: VTR rank list of top 10 Italian universities for UDA mathematics and computer science: E, G, A and L indicate numbers of outputs rated by VTR as excellent, good, acceptable, limited.

2.2 Bibliometric evaluation

Thomson Reuters' WoS was mined for raw data on the publications (articles, reviews, and conference papers) included in the Italian National Country Report. This raw data was then utilized for bibliometric analysis. Each article is then credited to the university scientist(s) who generated it utilizing this information and a complicated algorithm for reconciling the authors' affiliation and disambiguating the genuine author identities (D'Angelo et al., 2011).

The eight UDAs7 were evaluated because they are the types most likely to have research findings published in scientific publications, making bibliometric methods suitable for gauging performance. For increased relevance, we restricted our analysis to scientific discovery systems (SDSs) in which at least half of the scientists on staff published at least one paper between 2001 and 2003. About two-thirds of the total academic research workforce are SDSs (177 out of 205 in the hard sciences). These 177 SDSs averaged 26,241 scientists across 66 institutions during the study period (Table 2).

UDA	SDSs	Universities	Research staff	Publications
Agricultural and veterinary sciences	28	36	1,910	4,175
Biology	19	58	4,340	14,414
Chemistry	12	57	3,065	13,017
Earth sciences	12	47	856	2,131
Industrial and information engineering	42	59	3,596	13,867
Mathematics and computer sciences	9	56	2,197	6,384
Medicine	47	51	7,925	24,152
Physics	8	56	2,352	12,358
Total	177	66[8]	26,241	78,782*

Table 2**: **Research staff, publications and number of SDSs per UDA; data 2001-2003.

[7] Mathematics and computer sciences; physics; chemistry; earth sciences; biology; medicine; agricultural and veterinary sciences; industrial and information engineering.
[8] To make ranking lists comparable, the dataset has been limited to the 61 universities ranked by SCImago.

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

Indo-American Journal of Agricultural and Veterinary Sciences

Individual researchers, SDSs, UDAs, and institutions are all quantifiable units of research performance measurement. Since the VTR only gives rankings at the university and UDA levels, we have chosen to limit our analysis to those two overarching categories as well. However, bibliometric comparison of research institution performance should begin at the field level, i.e. SDS, because to the varying publishing intensities among scientific disciplines. Therefore, the SDS serves as our primary analytical unit. We use a labor productivity index called P to evaluate research output. Every institution participating in the SDS applies productivity measures to its research team. The MIUR9 has a database with information on the employees of all Italian universities, including their SDS categories. When measuring productivity, we focus on real results (or "impact") rather than just production (or "input"). The amount of citations received by the researcher's published works is used as a surrogate for actual results. However, as Abramo et al. (2008) pointed out, some scholars working in the same SDS publish in many WoS topic areas. The variable citation behavior of distinct topic categories necessitates a field (subject category) normalization when comparing the influence of academics working within the same SDS. The median number of citations10 for all Italian publications in the same year and WoS topic category11 is used to normalize a publication's citations. Since collaborative research is on the rise, we extrapolate each individual's true contribution to the final outcome. A university's SDS receives a proportional share of the total number of citations based on the number of co-authors employed by that SDS. According to their position on the list and the nature of their co-authorship (intra-mural or extra-mural), co-authors get varying weights for publications in the so-called "life science" categories (equivalent to 66 SDSs of 177).12. The last step in calculating a yearly performance average involves dividing each scientist's total number of years spent on staff at Italian institutions by the number of years included by the study. The assumption that universities' non-labor production components are dispersed uniformly is not far from the reality of Italy's publicly funded university system.13

The equations for the productivity P of university i in the SDS are as follows at the SDS level:

$$P = \frac{1}{\;} \cdot \sum^{N} \quad C_j \quad \cdot n[2]_{,s} \frac{\;}{m} \quad RS_{i,j=1} \quad j \quad j,i,s$$

[9]http://cercauniversita.cineca.it/php5/docenti/cerca.php, last accessed on Sept. 5, 2012.

[10]Observed as of 30/06/2009.

[11] For publications in multidisciplinary journals the standardized value is calculated as a weighted average of the standardized values for each subject category.

[12]For the life sciences, position in the list of authors reflects varying contribution to the work. Italian scientists active in these fields have proposed an algorithm for quantification: if the first and last authors belong to the same university, 40% of citations are attributed to each of them; the remaining 20% are divided among all other authors. If the first two and last two authors belong to different universities, 30% of citations are attributed to first and last authors; 15% of citations are attributed to second and last author but one; the remaining 10%are divided among all others. This algorithm could also be adapted to suit other national contexts.

[13] Prior to the VTR, all universities were almost completely financed through non-competitive MIUR allocation.

With:

$n_{j,i,s}$= fraction of authors of university *i* in SDS *s* on total co-authors of

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

publication j, (considering, if publication j falls in life science subject categories, the position of eachauthor in the list and the character of the co-authorship, intra-mural or extra-mural).

N_s = total number of publications in SDS s

$RS_{i,}$= research staff time equivalent of university i in SDS s, in the observed period

C_j = number of citations received by publication j

C_j^m= median of citations received by all Italian publications of the same year and subject category of publication j

To calculate the performance at UDA level, we aggregate productivity data of each SDS within the UDA. To account for: i) varying publication and citation intensities of different SDSs and ii) differing representativity, in terms of research staff, of the SDSs present in each UDA, data are conveniently: i) standardized and ii) weighted. At UDA level, in formula, productivity P of universit i in the UDA u is: $_{i,s}$

With:

$RS_{i,}$= research staff time equivalent of university i in UDA u, in the observed period

N_u = number of SDSs in the UDA u

$\bar{P}_s$= average value of productivity in SDS s of all universities

We apply the same procedure to calculate productivity of the entire university, againbeginning from the productivity of each SDS. In formula, productivity P per university i is: $=$ ^{1} $\cdot \sum P_{i,s} \cdot RS$—[4]

i $RS_i s{=}1\,\bar{P}_s$

i,s

With:

RS_i= research staff time equivalent of university i in the observed period

N_u = number of SDSs where the university is active

$\bar{P}_s$= average value of productivity in SDS s of all universities

2.3 Evaluation based on geographic and economic indicators

To generate zero-cost rankings to compare with those from the VTR and bibliometrics, we draw on economic and geographic information concerning the universities and their base regions. Intentionally, the rankings produced here are not based on the quality of research output.

The literature includes studies that show the relative performance of the various European regions in terms of articles cited. For example Bornmann and Leydesdorff (2011) show that, in Italy, the cities with highly cited articles (top 10%) in information science are particularly concentrated in the north of the nation. These results generated the idea of producing a ranking list of Italian universities in decreasing order of the numeric latitude of each home city.

Various studies also show the link between R&D expenditures and development at the regional level. Taking particular note of Guisan (2005), on the positive effects of university research expenditures on regional development, we have ranked

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

universities by university research expenditure per inhabitant in the relative home regions. We produce a third ranking list on the basis of GDP per inhabitant in the university home regions. For both of these lists we draw on data available from the Italian National Institute for Statistics[14] (ISTAT) concerning the year 2002, the middle year of the triennium under study 2001- 2003.

2.4 Evaluations available on Internet for free

Besides exploring the above indirect indicators of performance, we also examine the possibilities for using the yearly international university "league tables" (e.g. QS World University Rankings, Times Higher Education World University Rankings; The Leiden Rankings; ARWU; SCImago, etc). The most extensive free list is SCImago Institutions Ranking (SIR) World Report, which classifies 2500 universities per year[15] based on publication data from Scopus (Elsevier), with the results on Internet. Since the reports available do not extend as far back as 2003 we refer to the one closest to our period of observation, based on research output from 2004-2008. As justification, we verify that the ranking lists for bibliometric performance P for the period 2001-2003 (see Section 2.2) is strongly correlated (Spearman coefficient: +0.867, p-value= 0.000) to the P-based ranking for 2004-2008. From the SIR web site, we extract the ranks of Italian universities accordingto the indicators IC, Q1 and NI, where:

- IC, or International Collaboration, represents the percentage of publications co-authored with foreign organizations on total publications of the university.
- Q1, represents the ratio of scientific publications that a university manages to publish in the 25% of the most influential journals according to the SCImagoranking.
- NI, or Normalized Impact, represents the overall scientific impact of institutions. "Normalized Impact values show the ratio between the average scientific impact of an institution and the world average impact of publications of the same time frame, document type and subject area. Normalized Impact is computed using the methodology established by the Karolinska Intitutet in Sweden where it is named "*item oriented field normalized citation score average*"[16].

Finally, we use the reports on individual Italian university faculties published by CENSIS (Center for Social Investments Studies). These offer rankings by various performance indicators, from which we choose the list for the "Research" indicator, referring to competitive funding received by universities from the MIUR's PRIN programme for research projects of national interest. The reference VTR and P rankings are by UDA, therefore we take the CENSIS rankings of the individual faculties and aggregate them into the corresponding UDAs (Table 3).

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

Indo-American Journal of Agricultural and Veterinary Sciences

Faculty	UDAs
Engineering	Industrial and Information engineering
Medicine	Medicine
Agricultural sciences	Agricultural and Veterinary Sciences
Veterinary sciences	Agricultural and Veterinary Sciences
	Biology
	Chemistry
Natural and formal sciences	Physics
	Earth Science
	Mathematics and computer sciences

Table 3: Relation of UDAs to faculties

2. Results and analysis

This section presents the comparisons of the VTR and "zero-cost" rankings to the benchmark P ranking. Table 4 provides a brief review of all the rankings lists with their acronyms and sources and levels of data agglomeration

.

Acronym	Ranking by	Source of data	Level of analysis
VTR	VTR peer-review rating	CIVR	UDA/University
P	Biliometric rating	Web of Science	UDA/University
LAT	Latitude of university	Google maps	University
EXP	Research expenditure by universities (of a region)per inhabitant	ISTAT	University
GDP	Gross Domestic Product (of a region) per capita	ISTAT	University
IC	Rate of International Collaboration (%)	SCImago	University
Q1	Publications ratio in top journals (%)	SCImago	University
NI	Normalized impact	SCImago	University
RES	Funding received from PRIN program	CENSIS	Merged UDAs

Table 4: Acronyms of rankings used in the analysis

3.1. Rank correlations

To measure the level of accuracy, we calculate the Spearman rank correlations for each ranking. The report is divided in subsections, the first concerning the geographic and economic rankings, second the SCImago rankings, and finally the CENSIS ranking.

3.1.1 Geographic and economic rankings

Here we present the results from the correlation analyses relative to university bibliometric rank calculated on the basis of P, the VTR rankings, and rank from geographicand economic information, i.e. latitude (LAT), regional expenditures on R&D (EXP) and gross domestic product per capita (GDP).

We first carry out the correlations without distinguishing by UDA, since not all the rankings provide this level of

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

detail.

From Table 5, we see that there is a significant level of correlation between P rankings and LAT rankings (ρ=+0.6291), confirming the strong positive correlation[17] between university performance and latitude of the home city. These values are actually greater than the ones between P and VTR rankings (+0.6162).

The correlations to the two economic variables, EXP and GDP, are weaker, but still significant, at approximately +0.29 and +0.35.

	P	VTR	LAT	EXP	GDP
P	1.0000				
VTR	0.6162*	1.0000			
LAT	0.6291*	0.5903*	1.0000		
EXP	0.2904*	0.1289	0.0584	1.0000	
GDP	0.3511*	0.1639	0.3549*	0.1120	1.0000

Table 5: Spearman correlation matrix among P, VTR and geographic and economic university rankings
** p-value< 0.05*

Focusing only on the ranking by variable with greatest correlation, i.e. latitude, we carryout further analyses at the UDA level. The results (Table 6) show a certain variability across UDAs and offer two points of reflection. First, the correlation values per UDA are all lower than the overall value per university; second, correlation still remains strong for the biomedical area. The highest value of correlation between LAT and P is seen in Medicine (ρ =+0.6081). Also, for four UDAs out of eight, the value of correlation between P and LAT is significantly higher than that between P rank and VTR rank by peer-review. For example in Agricultural and veterinary sciences the value of correlation for LAT to P is +0.4974, compared to +0.4033 for VTR to P. In earth sciences, where the correlation between ranks by VTR and P is not significant (ρ =+0.1771), the correlation for LAT to P ranks is noticeably higher, and significant (ρ =+0.4642). The correlations for the remaining cases are quite similar except for Industrial and information engineering and Chemistry, where correlation between the P and VTR ranks is decidedly higher (ρ =+0.4920).

[17] We followed the guidelines by Cohen (1988) on the strength of association.

	Agricultural and veterinary Sciences				*Biology*		
	P	VTR	LAT		P	VTR	LAT
P	1.0000			P	1.0000		
VTR	0.4033*	1.0000		VTR	0.5267*	1.0000	
LAT	0.4974*	0.3725*	1.0000	LAT	0.4646*	0.3908*	1.0000
	Chemistry				*Earth Sciences*		
	P	VTR	LAT		P	VTR	LAT
P	1.0000			P	1.0000		
VTR	0.4920*	1.0000		VTR	0.1771	1.0000	
LAT	0.3063*	0.2601	1.0000	LAT	0.4642*	0.1213	1.0000
	Industrial and information engineering				*Mathematics and computer sciences*		
	P	VTR	LAT		P	VTR	LAT
P	1.0000			P	1.0000		
VTR	0.4250*	1.0000		VTR	0.4486*	1.0000	

Indo-American Journal of Agricultural and Veterinary Sciences

LAT	0.4279*	0.4822*	1.0000	LAT	0.3241*	0.3885*	1.0000
	Medicine				*Physics*		
	P	VTR	LAT		P	VTR	LAT
P	1.0000			P	1.0000		
VTR	0.4663*	1.0000		VTR	-0.0389	1.0000	
LAT	0.6081*	0.4125*	1.0000	LAT	0.1469	0.2578	1.0000

Table 6: Spearman correlation matrix among P, VTR and geographical ranking by UDA
** p-value< 0.05.*

3.1.2 SCImago ranking

This second section presents analysis of correlations with ranks derived from the SCImago rankings. In Table 7 we see that the NI ranking is highly correlated with P ranking (ρ=+0.7225) and the VTR and P rankings are also correlated but with lower intensity (ρ=+0.6162). The correlation between the IC and P ranking is also strong but with still lower intensity (ρ=+0.5219).

Given the results showing higher correlation for the NI indicator, we decided to carry out a more detailed analysis of the relationship at the UDA level. Again in this case, since SCImago does not provide rankings by area, we assign the same value of normalized impact to every UDA in each ranked university. The results continue to be significant, with very high correlation values for all UDAs. The highest value is obtained in Agricultural and veterinary sciences (ρ=+0.6623), much above the correlation obtained with VTR rank (ρ=+0.4033). For the Earth sciences area the correlation with NI is +0.4142, while correlation with VTR rank is not significant. For the Physics area the correlation is not significant for both NI and VTR (Table 8).

	P	VTR	IC	Q1	NI
P	1.0000				
VTR	0.6162*	1.0000			
IC	0.5219*	0.5889*	1.0000		
Q1	0.2572*	0.5026*	0.3367*	1.0000	
NI	0.7225*	0.6691*	0.5210*	0.3973*	1.0000

Table 7: Spearman correlation matrix among P, VTR and SCImago rankings

	Agricultural and veterinary Sciences				*Biology*		
	P	VTR	NI		P	VTR	NI
P	1.0000			P	1.0000		
VTR	0.4033*	1.0000		VTR	0.5267*	1.0000	
NI	0.6623*	0.4557*	1.0000	NI	0.4981*	0.5041*	1.0000
	Chemistry				*Earth Sciences*		
	P	VTR	NI		P	VTR	NI
P	1.0000			P	1.0000		
VTR	0.4920*	1.0000		VTR	0.1771	1.0000	
NI	0.3986*	0.2234	1.0000	NI	0.4142*	0.1005	1.0000
	Industrial and information engineering				*Mathematics and computer sciences*		

	P	VTR	NI		P	VTR	NI
P	1.0000			P	1.0000		
VTR	0.4250*	1.0000		VTR	0.4486*	1.0000	
NI	0.4800*	0.5104*	1.0000	NI	0.3935*	0.5230*	1.0000
		Medicine				*Physics*	
	P	VTR	NI		P	VTR	NI
P	1.0000			P	1.0000		
VTR	0.4663*	1.0000		VTR	-0.0389	1.0000	
NI	0.5483*	0.4921*	1.0000	NI	0.0784	0.1832	1.0000

Table 8: Spearman correlation matrix among P, VTR and NI SCImago ranking, by UDA

3.1.3 CENSIS ranking

This final subsection concerns the analysis of correlation between P rank and the ranks derived from CENSIS. In this case the ranks are presented by macro-UDA, defined in Table 3, since CENSIS ranks are given only by faculty.

The biomedical area, in this case included in the Natural and formal sciences macro- UDAs, continues to show high correlation values (Table 9). For this macro-UDA the correlation between indicator RES and P is actually strong (ρ=+0.6342), greater than that between P and VTR (ρ=+0.5930). Among the remaining areas, for Engineering and agricultural and veterinary sciences the correlation with the CENSIS ranking is low and notsignificant.

	Agricultural and veterinary sciences				*Natural and formal sciences*		
	P	VTR	RES		P	VTR	RES
P	1.0000			P	1.0000		
VTR	0.4060*	1.0000		VTR	0.5930*	1.0000	
RES	0.1717	0.3608	1.0000	RES	0.6342*	0.5043*	1.0000
		Medicine				*Engineering*	
	P	VTR	RES		P	VTR	RES
P	1.0000			P	1.0000		
VTR	0.5527*	1.0000		VTR	0.4553*	1.0000	
RES	0.5819*	0.4921*	1.0000	RES	0.2871	0.0050	1.0000

Table 9: Spearman correlation matrix among P, VTR and CENSIS ranking, by macro-UDA

3.2 Distributions of changes in rank

The results of the Spearman correlation analyses evidence strong association between the P rankings and those derived from informed rankings, particularly from SCImago's NI normalized impact value. However a correlation value around 0.6 does not provide any specific information on the behavior of specific categories of universities, such as those classified at the top of the ranking. Thus we classify the institutions into four classes, as is common in research assessment exercises, according to their percentile ranking. We assign values of 4, 3, 2 and 1, corresponding to the first, second, third and fourth quartiles for the performance value. We then

Indo-American Journal of Agricultural and Veterinary Sciences

calculate the distribution of the shifts in quartile when a university "shifts" from the P ranking to SCImago NI and peer-review VTR ranking. Table 10 shows that, in comparing the NI ranking to the benchmark bibliometric P, 54.10% of universities remain in the same performance quartile[18]; 32.79% move one quartile and 13.11% make a two quartile jump. No university makes a three quartile jump. However, comparing the VTR and P ranking, the distribution of shifts is slightly different, showing a longer tail to the right: less universities maintain a constant quartile (45.90%) and a greater number shift one quartile (40.98%); 1.64% of the universities actually make the maximum shift of three quartiles.

	NI vs P	NI vs VTR	VTR vs P
Changes	*Relative frequency distributions*		
0	54.10%	47.54%	45.90%
1	32.79%	39.34%	40.98%
2	13.11%	13.11%	11.48%
3	0.00%	0.00%	1.64%
	Cumulative frequency distributions		
≤ 0	54.10%	47.54%	45.90%
≤ 1	86.89%	86.89%	86.89%
≤ 2	100.00%	100.00%	98.36%
≤ 3	100.00%	100.00%	100.00%

Table 10: Distributions of change in rank among NI, VTR and P.

[18] We note that the value change of rank within quartiles for any universities which do not shift quartile, maybe larger than that of universities that shift quartile.

This first analysis of the quartile shift distributions indicates that the VTR is less reliable than the NI ranking. Still, we would like more detailed examination of the case of the universities ranked as "top" by P ranking. Table 11 presents various simulations for extraction of top universities for the first two quartiles of the P ranking, with calculation of the numbers of universities that would not register as such under NI or VTR rank. For example, if we consider the top 25% of universities, i.e. those that fall in performance quartile 4, we observe that 7 universities out of 15 (46.7%) are not classified as such for theVTR. This number falls to 4 out of 15 (26.7%) if we consider the ranking for NI.

In general, for all simulations of extracting top universities that we try, from top 5% to top 50%, we observe that with the exception of three cases, the NI rank is always closer than the VTR to the more trustworthy P rank. Only in the case of the extracting the top 5% by P do we obtain superimposition of the rank from VTR (0 variations out of three universities), but given the very slim number of cases that compose this group the data has little relevance.

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013

Indo-American Journal of Agricultural and Veterinary Sciences

	NI vs P		VTR vs P	
Top universities	Variations	Percentage	Variations	Percentage
5%	2 out of 3	66.67%	0 out of 3	0.00%
10%	2 out of 6	33.33%	3 out of 6	50.00%
15%	2 out of 9	22.22%	5 out of 9	55.56%
20%	4 out of 12	33.33%	5 out of 12	41.67%
25% (4th quartile)	4 out of 15	26.67%	7 out of 15	46.67%
30%	4 out of 18	22.22%	8 out of 18	44.44%
35%	6 out of 21	28.57%	8 out of 21	38.10%
40%	6 out of 24	25.00%	8 out of 24	33.33%
45%	9 out of 27	33.33%	7 out of 27	25.93%
50% (3rd quartile)	9 out of 30	30.00%	7 out of 30	23.33%

Table 11: Percentile change in rank among NI, VTR and P

3. Conclusions

In previous works the authors had demonstrated the limits in planning of the Italian national research assessment exercise and, for the hard sciences, the undisputable superiority of bibliometrics over the peer-review methodology adopted. In this work we have asked whether it would be possible to prepare university ranking lists at zero cost and with levels of accuracy comparable or superior to those of the VTR ranking list.

Taking a ranking of Italian universities by decreasing latitude from north to south, we found it comparable to the VTR: for the individual disciplines, the results actually showed greater accuracy than the VTR in half the disciplines and lesser accuracy in three out of eight. Again at the level of discipline, the freely-available CENSIS rankings were also equivalent to those from the VTR.

However when we compared SCImago university rankings by average citation impact we found that these lists outperformed the VTR, both at the absolute level and by discipline.

The moral: not only would the application of bibliometric techniques be more precise, more robust and notably less expensive, but the entire direct and indirect costs of the VTR

for the hard sciences could be completely avoided by resorting to zero-cost rankings, and these would give results of equal or greater level of accuracy.

Governments in general, and especially the Italian government, should question the competencies of those who are planning national evaluation exercises, or at least ensure that there is a sufficient exchange of knowledge between scholars and practitioners to ensure maximum efficiency, effectiveness and fairness in their conduct.

References

Abramo, G., D'Angelo, C.A., Di Costa, F. (2011). National research assessment exercises: a comparison of peer-review and bibliometrics rankings. *Scientometrics*, 89(3), 929- 941.

Abramo, G., D'Angelo, C.A. (2011). Evaluating research: from informed peer-review to bibliometrics, *Scientometrics*, 87(3), 499-514.

Abramo, G., D'Angelo, C.A., Di Costa, F. (2008). Assessment of sectoral aggregation distortion in research productivity measurements. *Research Evaluation*, 17(2), 111–121.

Aksnes, D.W., Taxt, R.E. (2004). Peer reviews and bibliometric indicators: a comparative study at Norvegian University. *Research Evaluation*, 13 (1), 33–41.

Bornmann, L., Leydesdorff, L. (2012). Which are the best performing regions in information science in terms of highly cited papers? Some improvements of our previous mapping approaches. *Journal of Informetrics*, 6(2), 336-345.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, Second Edition, Lawrence Erlbaum Associates, Hillsdale, NJ.

D'Angelo, C. A., Giuffrida, C., Abramo, G. (2011). A heuristic approach to author name disambiguation in large-scale bibliometric databases. *Journal of the American Society for Information Science and Technology*, 62(2), 257–269.

ERA (2010). The Excellence in Research for Australia Initiative http://www.arc.gov.au/era/(last accessed on Sept. 5, 2012).

Franceschet, M., Costantini, A. (2011). The first Italian research assessment exercise: A bibliometric perspective. *Journal of Informetrics*, 5(2), 275–291.

Guisan, M.C. (2005). Universities and research expenditure in Europe and the Usa, 1993- 2003: an analysis of countries and regions. *Regional and Sectoral Economic Studies, AEEADE*, 5(2).

Horrobin, D.F. (1990). The philosophical basis of peer-review and the suppression of innovation. *Journal of the American Medical Association*, 263(10), pp. 1438-1441.

MacRoberts, M.H., MacRoberts, B.R. (1996). Problems of citation analysis. *Scientometrics*, 36(3), 435–444.

Meho, L.I., Sonnenwald, D.H. (2000). Citation ranking versus peer evaluation of senior faculty research performance: a case study of Kurdish Scholarship. *Journal of the American Society for Information Science*, 51(2), 123-138.

Moed, H.F. (2002). The impact-factors debate: the ISI's uses and limits. *Nature*, 415, 731– 732.

Moxam, H., Anderson, J. (1992). Peer-review. A view from the inside. *Science and Technology Policy*, pp. 7-15.

Oppenheim, C. (1997). The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology. *Journal of Documentation*, 53, 477–487.

Oppenheim, C., Norris, M. (2003). Citation counts and the research assessment exercise V: archaeology and the 2001 RAE. *Journal of Documentation*, 56 (6), 709–730.

Pendlebury, D.A. (2009). The use and misuse of journal metrics and other citation indicators. *Scientometrics*, 57(1), pp. 1-11.

RAE (2008). Research Assessment Exercise. http://www.rae.ac.uk/aboutus/(last accessedon Sept. 5, 2012).

Rinia, E.J., van Leeuwen, Th.N., van Vuren, H.G., van Raan, A.F.J. (1998). Comparative analysis of a set of bibliometric indicators and central peer-review criteria, Evaluation of condensed matter physics in the Netherlands. *Research Policy*, 27(1), 95–107.

Serenko, A., Dohan, M. (2011).

ISSN 2321–9602 www.iajavs.com
Vol. 1, Issue 1 Feb 2013
Indo-American Journal of Agricultural and Veterinary Sciences

Comparing the expert survey and citation impact journal ranking methods: Example from the field of Artificial Intelligence. *Journal of Informetrics*, 5(4), 629-648.

Thomas, P.R., Watkins, D.S. (1998). Institutional research rankings via bibliometric analysis and direct peer-review: A comparative case study with policy implications. *Scientometrics*, 41(3), 335-355.

Van Raan, A.F.J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), pp. 133-143.

Van Raan, A.F.J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, 67(3), 491–502.

VTR (2006). *Three-year research evaluation (2001–2003). Results separated by scientific area*. http://vtr2006.cineca.it/index_EN.html, Last accessed on Sept. 5, 2012

.