Check for updates

RESEARCH ARTICLE

Journal ratings as predictors of articles quality in Arts, Humanities and Social Sciences: an analysis based on the Italian Research Evaluation Exercise [version 1; peer review: 3 approved]

Andrea Bonaccorsi, Tindaro Cicero, Antonio Ferrara, Marco Malgarini

ANVUR, Via Ippolito Nievo 35, Rome, 00153, Italy

V1 First published: 07 Jul 2015, 4:196
<https://doi.org/10.12688/f1000research.6478.1>Latest published: 07 Jul 2015, 4:196
<https://doi.org/10.12688/f1000research.6478.1>

Abstract

The aim of this paper is to understand whether the probability of receiving positive peer reviews is influenced by having published in an independently assessed, high-ranking journal: we eventually interpret a positive relationship among peer evaluation and journal ranking as evidence that journal ratings are good predictors of article quality. The analysis is based on a large dataset of over 11,500 research articles published in Italy in the period 2004-2010 in the areas of architecture, arts and humanities, history and philosophy, law, sociology and political sciences. These articles received a score by a large number of externally appointed referees in the context of the Italian research assessment exercise (VQR); similarly, journal scores were assigned in a panel-based independent assessment, which involved all academic journals in which Italian scholars have published, carried out under a different procedure. The score of an article is compared with that of the journal it is published in: more specifically, we first estimate an ordered probit model, assessing the probability for a paper of receiving a higher score, the higher the score of the journal; in a second step, we concentrate on the top papers, evaluating the probability of a paper receiving an excellent score having been published in a top-rated journal. In doing so, we control for a number of characteristics of the paper and its author, including the language of publication, the scientific field and its size, the age of the author and the academic status. We add to the literature on journal classification by providing for the first time a large scale test of the robustness of expert-based classification.

Keywords

Journal rankings, ANVUR, Arts, Humanities and Social Sciences, Journal ratings, Article Quality

Open Peer Review

Approval Status

	1	2	3
version 1			
07 Jul 2015	view	view	view

1. **Geoffrey Williams**, EvalHum Initiative and
Université de Bretagne-Sud, Lorient, France

2. **Alesia Zuccala**, University of Copenhagen,
Copenhagen, Denmark

3. **Chiara Faggiolani**, Università degli Studi di
Roma, Rome, Italy

Domenica Fioredistella Iezzi, Università
degli Studi di Roma Tor Vergata, Rome, Italy

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the **Research on Research, Policy & Culture gateway**.

This article is included in the **Proceedings of the 2015 ORCID-Casrai Joint Conference collection**.

Corresponding author: Marco Malgarini (marco.malgarini@anvur.it)

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2015 Bonaccorsi A *et al.* This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the **Creative Commons Zero "No rights reserved" data waiver** (CC0 1.0 Public domain dedication).

How to cite this article: Bonaccorsi A, Cicero T, Ferrara A and Malgarini M. **Journal ratings as predictors of articles quality in Arts, Humanities and Social Sciences: an analysis based on the Italian Research Evaluation Exercise [version 1; peer review: 3 approved]** F1000Research 2015, 4:196 <https://doi.org/10.12688/f1000research.6478.1>

First published: 07 Jul 2015, 4:196 <https://doi.org/10.12688/f1000research.6478.1>

Introduction

There is a large degree of agreement on the notion that research assessment in humanities and social sciences (HSS) is made more complex by a variety of factors. First, in these fields the structure of academic publication is largely different, with a large weight assigned to books and monographs, and to production in national language (Finkenstaedt, 1990). Consequently, the bibliometric approach is considered to be of limited usefulness (Nederhof *et al.*, 1989), not only because journals are a small fraction of total production and indexed journals are a tiny fraction of the population of journals in HSS, but also because the meaning of citations is different (Frost, 1979). Third, and even more challenging, there is evidence that the number of research quality criteria is larger in HSS than in other fields, and also that there is less agreement on these criteria (Hemlin, 1996; Hemlin & Gustafsson, 1996; Hug *et al.*, 2013; Hug *et al.*, 2014; Ochsner *et al.*, 2012; Ochsner *et al.*, 2013).

Faced with these challenges, the state of the art of the assessment of research in HSS at international level has followed several directions. On the one hand, it is agreed that peer review is still the most important evaluation methodology, so large efforts are made in making it more sophisticated, methodologically controlled, based on sound principles of evaluation methodology in social sciences, and free from unwanted biases, distortions and unexpected side effects. Under this agenda, issues such as the notion of originality, unorthodox science, or interdisciplinarity are under examination (Guetzkow *et al.*, 2004; Hammarfelt, 2011). On the other hand, there are many efforts to classify and evaluate non-indexed journals (mainly in national languages), as one of the main vehicles for academic communication. An additional line of work refers to the classification of books and publishers.

This paper reports on a large experiment in the classification of journals in HSS carried out in Italy in the 2012–2014 period for the National Scientific Habilitation (Abilitazione Scientifica Nazionale; ASN). The exercise was based upon a mandatory provision in the law to rate *all* journals, in order to calculate the overall academic production of all candidates to the national procedure to become associate professor or full professor. This exercise asked the National Agency for the Evaluation of Universities and Research Institutes (ANVUR) to evaluate all journals in which at least one Italian scholar published at least one paper in the 2002–2012 period, for a total of more than 60,000 titles.

While the rating of journals has been followed in several national contexts, it is only in the Italian exercise that there is the opportunity to carry out a controlled experiment in order to test the robustness of journal classification. In fact, we have two independent evaluations carried out on the same set of journals. On the one hand, a panel of experts classified all journals as academic and non-academic (i.e. popular, professional, technical, cultural and political etc.), and rated the subset of academic journals in A-rated and non A-rated. The rating exercise was done on the basis of the reputation, esteem, diffusion and impact of journals, that is, on a qualitative, expert-based, reputational basis. On the other hand, we also have the rating of individual articles published in those journals, which have been done by a large number of individual referees (not panels) and summarized with a consensus agreement approach by expert panels,

who however acted independently from the other panels, and without exchange of information. This peculiarity of the Italian context and the time sequence of events creates a favorable condition for carrying out a controlled experiment.

This paper extends to all HSS, with the exception of economics and business, the analysis initiated by Ferrara & Bonaccorsi (2015) on journals in the area of philosophy and history. In the following, we first introduce the database used for the analysis and hence we test for the influence of the journal class on the article score. Some consideration on the results obtained will conclude the paper.

Methods

The paper is based on a dataset including data on all the journal articles submitted for evaluation by Italian scholars in the disciplinary areas of architecture, arts and humanities, history and philosophy, law and sociology and political science. Submissions for evaluation took place within the framework of VQR 2004–10, Italy's national research assessment exercise involving all professors and researchers affiliated to the Italian universities and Public Research Organizations (PROs) as of November 2011. According to adopted rules, research evaluation in HSS was entirely based on peer review; research quality was assessed against the criteria of *relevance*, meaning contribution to the advancement of the state of art in the field, also in terms of adequacy, efficacy, timeliness and duration of impacts; *originality and innovation*, meaning contribution to the creation of new knowledge in the field; *internationalization*, meaning position in the international research landscape. Evaluation has been conducted by five Groups of Evaluation Experts (GEV in the Italian acronym), one for each area in HSS (Architecture; Arts and Humanities; History and Philosophy; Law; Sociology and political sciences); reviewers were instructed by GEV to evaluate articles only on the basis of their merit, regardless of the journal in which they are published in and of the language of publication. Each article had a possible rating of Excellent (A), Good (B), Fair (C) or Limited (D); to each class corresponded a score ranging from 1 (for articles A-rated) to zero (for articles deemed as limited). Negative scores were also assigned in case the article was deemed as non-academic (-1) or for plagiarism or fraud (-2, see Ancaiani *et al.*, 2015 for details). Limited to the human and social sciences, a substantial fraction of articles – namely, 6,701 out of 11,660 (Table 1) – appeared on journals deemed as 'A-class' according to the procedure of ASN, intended to select the best researchers for the ranks of associate and full professors. Those journals, according to the relevant Ministerial Decree (No. 76/2012), were those 'internationally recognized as excellent because of the rigor of their procedures of peer review and because of their diffusion among, esteem by, and impact on, the scholarly community of a field, as indicated also by their presence in the major national and international databases' (our translation). Most of the remaining articles appeared on journals deemed as 'academic' for the purposes of the ASN, while a minority were published in journals that remained 'uncategorized'. The main feature of the dataset, thus, is that it allows the comparison between the evaluations of journals and individual articles.

A preliminary analysis shows that there is a relationship between the evaluation of individual articles and that of journals where the article is published (Table 2). The non-parametric test for categorical

Table 1. Description of dataset.

Area of assessment	Acronym	Full Professor	Associate Professor	Researcher	Other	N° of articles	N° of articles in Class A journals
Architecture	Area08	280	278	353	7	918	360
Antiquities, philology, literary studies, art history	Area10	1,040	1,191	1,322	26	3,579	1,954
History, philosophy, pedagogy and psychology	Area11	713	680	726	8	2,127	1,086
Law	Area12	1,488	983	1,337	30	3,838	2,637
Political and social sciences	Area14	338	409	442	9	1,198	664
Total		3,859	3,541	4,180	80	11,660	6,701

Table 2. Preliminary analysis of association between the evaluation of research product and the evaluation of journal.

		Evaluation of journal			
		A	Not A	Not academic	Total
Evaluation of research product	A	1,344	573	20	1,937
	B	3,184	1,743	92	5,019
	C	1,322	1,096	80	2,498
	D	837	1,176	150	2,163
	Non-academic and other	14	21	8	43
Total		6,701	4,609	350	11,660

Pearson $\chi^2=630.9$; $p\text{-value}=0.000$

data (Pearson χ^2) is statistically significant at 1% (All the statistical analyses have been performed using the software STATA ver. 13 (<http://www.stata.com/stata13/>)), showing that the two distributions are not independent and hence the two ratings are mutually related. In the following, we will analyze more thoroughly this relationship, also controlling for a number of author-level and article-level variables.

The influence of journal classification on the article score

We assume that the probability for an article i , published in the journal j , of receiving a score equal to $x \in \{-2; 1\}$ is influenced by the class assigned to the journal, once controlling for a number of characteristics of the article:

$$P(\text{Score}_{ij} = x) = F(\text{Journal class}_{ij}, \text{Paper characteristics}_{ij}) \quad (1)$$

Among the controls, we consider the language of publication (Italian or not) and the age (distinguishing among 3 age classes, less than 40 years, between 41 and 55 years and more than 55 years), scientific sector of activity (Scientific Areas 8, 10, 11, 12, 14), academic status (full professor; associate professor; researcher; other) and gender of the researcher. We also add the consideration of two binary variables controlling for the existence of international

co-author(s) and for the nationality of the referees (allowing for the possibility of international referees). We finally add a variable taking into account the size of the scientific area of the author. The model is estimated as an ordered probit, an extension of the standard binary probit model, used when the dependent variable takes the form of a ranked and multiple discrete variable, considering alternatively the whole sample or each scientific area; in the first case, we also control for possible area-specific effects. In order to avoid the “dummy trap”, we normalize with respect to articles written in Italian with no international co-author, evaluated by an Italian reviewer, presented by a female researcher in sociology and political science, aged less than 40: i.e. the statistical significance, sign and magnitude of estimated parameters are to be interpreted as differentials with respect to this control group. The total number of available observations amounts to 11,660 varying from a minimum of 918 in architecture to a maximum of 3,838 in law (Table 3).

The main result is that both at the aggregate level and in each scientific area the article score is higher as the journal ranking gets better: in other words, the probability of receiving a high score grows if the article is published in a high-ranking journal according to the evaluation of the ASN’s experts. As for the control variables, we confirm most of the results already emerged in a previous paper on the same data (Cicero *et al.*, 2014), namely, that article

Table 3. Ordered probit model (Dependent variable: article score).

Variables	Total	Architecture	Arts, & Hum.	Hist. & Phil.	Law	Sociology & Pol. Sci.
Journal rating	0.417***	0.542***	0.400***	0.379***	0.503***	0.323***
Architecture	0.134***					
Arts and Humanities	0.720***					
Hist. & Philosophy	0.471***					
Law	0.259***					
Italian language	-0.372***	-0.518***	-0.148***	-0.623***	-0.281***	-0.704***
41–55 years	-0.151***	-0.256	-0.208***	0.0492	-0.265***	-0.265**
More than 55 years	-0.582***	-0.662***	-0.726***	-0.394***	-0.563***	-0.572***
Associate professor	0.318***	0.206**	0.308***	0.265***	0.439***	0.234***
Full professor	0.818***	0.788***	0.690***	0.660***	1.096***	0.679***
Other personnel	-0.277**	-0.157	-0.329	0.359	-0.415**	-0.742*
Male	0.0777***	0.184**	0.0882**	0.0257	0.0542	-0.00491
International coauthors	0.301***	0.505***	0.122	0.237	0.902***	0.205
International reviewer	0.153***	0.363***	0.185***	0.0243	0.225***	0.0722
Number of Professor in the scientific sector (SSD)	-0.00131***	-0.00297***	-0.00248***	-0.00146**	-0.000921***	-0.00390***
Constant cut1	-2.854***	-2.514***	-2.081***	-2.990***	-1.867***	-2.953***
Constant cut2	-1.718***	0.363	-0.454***	-2.106***	-1.777***	-0.490***
Constant cut3	-1.695***	1.076***	0.200*	-0.357**	0.378***	0.402**
Constant cut4	0.302***	2.387***	1.574***	0.318**	1.162***	1.539***
Constant cut5	1.026***			1.651***	2.710***	
Constant cut6	2.400***					
Observations	11,660	918	3,579	2,127	3,838	1,198
Pseudo R-squared	0.0814	0.0919	0.0543	0.0720	0.0926	0.0832

*** p<0.01, ** p<0.05, * p<0.1

scores are higher for papers not written in Italian, with international co-authors, published by an under-40, male full or associate professor. Moreover, we also find that at aggregate level and in most areas an international reviewer and a lower number of professors in the specific scientific sector (SSD) are associated with an higher article score: a possible interpretation of the first result is that the expert groups responsible for the evaluation (GEV) mostly assign to international reviewers more internationalized papers, that are considered to have an higher probability of receiving a high score, given also that the level of internationalization was one of the evaluation criteria according to VQR rules (see again [Ancaiani et al., 2015](#)). As for a negative relationship among area size and article score, this result emerged already in [Ferrara & Bonaccorsi \(2015\)](#) for the scientific fields in history and philosophy and is now extended to all HSS: a possible interpretation is that small fields may be favored by a “proximity bias” among authors and reviewers, thus resulting, *ceteris paribus*, in higher article scores.

As a final check, we concentrate on the probability of receiving an excellent score and relate it to the fact that the article is published

in a top, A-Class journal, once controlling for the same variables considered in model 1:

$$P(\text{Score}_{ij} = \text{“E”}) = F(\text{Journal class}_{ij} = \text{“A”}, \text{Paper characteristics}_{ij}) \quad (2)$$

In (2), F is the logistic function and the model is estimated as a logit, a class of models allowing to predict the binary response based on the specified predictors. A desirable feature of the logit model is that the regression coefficients may easily be transformed in odds ratio, expressing the change in the odds of the occurrence under scrutiny (in our case, the odds for a paper of receiving an ‘Excellent’ evaluation) due to a small change of a given predictor: in our case, we are particularly interested in the odds associated with the classification of a journal as a top, Class A journal. Estimation results for both the aggregate sample and each scientific area are presented in [Table 4](#).

According to logit estimations, the probability of receiving an excellent evaluation is positively affected by the journal in which the paper is published in: more specifically, publishing in a class A

Table 4. Logit model (Odds ratio).

Variables	Total	Architecture	Arts, & Hum.	Hist. & Phil.	Law	Sociology & Pol. Sci.
Top Journal Classification	1.952***	2.513***	1.834***	2.424***	1.990***	1.311
Architecture	1.210					
Arts and Humanities	3.042***					
History and Philosophy	2.031***					
Law	1.084					
Italian language	0.488***	0.311***	0.681***	0.333***	0.529***	0.243***
41–55 years	0.878	0.408**	0.697**	1.144	0.807	0.671
More than 55 years	0.411***	0.248***	0.303***	0.506**	0.572***	0.252***
Associate professor	1.793***	1.283	1.815***	1.825***	2.629***	1.620*
Full professor	4.650***	3.263***	3.831***	4.023***	9.909***	4.877***
Other personnel	1.660	-	1.360	3.057	2.470	1.293
Male	1.247***	1.664**	1.263***	1.155	1.077	1.028
International coauthors	1.611***	2.357**	1.118	1.558	5.149***	1.511
International reviewer	1.352***	1.566**	1.393***	1.178	1.560***	1.490**
Full prof. in SSD	0.998**	0.992**	0.996**	0.998	1.000	0.990***
Constant	0.065***	0.201***	0.258***	0.137***	0.0332***	0.236***
Observations	11,660	911	3,579	2,127	3,838	1,198
Pseudo R-squared	0.116	0.140	0.0738	0.122	0.129	0.143

*** p<0.01, ** p<0.05, * p<0.1

journal almost doubles the probability of receiving an excellent evaluation. Looking at the results in each scientific area, the odds of receiving an excellent evaluation are more than doubled by the publication in a Class A journal in architecture and history and philosophy; the effect is somewhat lower, but still highly significant, in law, and arts and humanities, while disappearing in sociology and political sciences. Logit estimation also broadly confirms the results already emerging from the ordered probit model: the odds of receiving an excellent evaluation are increased by publishing in a foreign language, with an international co-author (albeit only in law and architecture) and when the submitting author is 40 years old or younger, an associate or full professor and a male. Gender effect is in fact significant at the aggregate level and in architecture and humanities, but not in the remaining areas. Also in this case, having an international reviewer and publishing in a SSD characterized by a lower number of full professors helps in obtaining an excellent evaluation.

Conclusions

Using a very large dataset of journal articles published in HSS, the paper proves that independent classifications of journals may be considered as good predictors of the score assigned to individual articles. More specifically, we find that, after controlling for a number of articles' characteristics, the probability of receiving a better score grows with the quality profile of the journal the article

is published in; moreover, the probability of receiving an excellent score almost doubles when the paper is published in a top, A-Class journal. The findings hold both at the aggregate level and for each specific sub-area considered in the analysis. While peer review has to remain the main evaluation methodology, our results indicate that journal classifications may be considered as a useful supporting tool in large evaluation exercise, since it may provide reviewers with valuable information apt to support expert evaluation.

Data availability

The authors hold the view that it is important to allow the free access to data used in the article in order to enable others to replicate the study. However, information used in the article were gathered by the national agency responsible for evaluation of the University and research system in Italy (ANVUR), in the framework of this VQR exercise. In this context, ANVUR asked Italian professors to provide access to their publications, assuming the commitment not to disclose to the public, unless in an aggregate form, any data concerning the publications submitted for the evaluation and, most importantly, the results of the evaluation itself. This is deemed as necessary in order to guarantee the full anonymity of evaluations performed on each individual publications and on each Italian professor. For this reason, as the public agency in charge of evaluating research of Italian universities, ANVUR does not allow to make information about individual evaluations available to the general public.

The information used to generate data in this article concerning journal classification is available to the public at the following URL: http://www.anvur.it/index.php?option=com_content&view=article&id=254&Itemid=315&lang=it.

Author contributions

The paper is the result of a common effort of the authors. However, Andrea Bonaccorsi can be credited for the “Introduction” and the “Conclusions”, while Antonio Ferrara took care of the “Methods”

section and Tindaro Cicero and Marco Malgarini were jointly responsible for the estimates contained in the “The influence of journal classification on the article score” section. All authors have read and agreed to the final content of the manuscript.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

References

Ancaiani A, Anfossi A, Barbara A, *et al.*: **Evaluating Scientific Research in Italy: the 2004–10 Research Evaluation Exercise.** *Res Eval.* 2015.

[Publisher Full Text](#)

Cicero T, Malgarini M, Benedetto S: **Research quality, characteristics of publications and socio-demographic features of Universities and Researchers: evidence from the Italian VQR 2004–2010 evaluation exercise.** *Proceedings of the science and technology indicators Conference 2014 Leiden “Context Counts: Pathways to Master Big and Little Data”.* 2014.

[Reference Source](#)

Ferrara A, Bonaccorsi A: **How robust is journal rating in Humanities and Social Sciences? Evidence from a large scale multi-method exercise.** Submitted for publication. 2015.

Finkenstaedt T: **Measuring research performance in the humanities.** *Scientometrics.* 1990; **19**(5–6): 409–417.

[Publisher Full Text](#)

Frost CO: **The use of citations in literary research: a preliminary classification of citation functions.** *The Library Quarterly: Information, Community, Policy.* 1979; **49**(4): 399–414.

[Reference Source](#)

Guetzkow J, Lamont M, Mallard G: **What is originality in the humanities and the social sciences?** *Am Sociol Rev.* 2004; **69**(2): 190–212.

[Publisher Full Text](#)

Hammarfelt B: **Interdisciplinarity and the intellectual base of literature studies: Citation analysis of highly cited monographs.** *Scientometrics.* 2011; **86**(3): 705–725.

[Publisher Full Text](#)

Hemlin S: **Social studies of the humanities: a case study of research**

conditions and performance in ancient history and classical archaeology, and English. *Res Eval.* 1996; **6**(1): 53–61.

[Publisher Full Text](#)

Hemlin S, Gustafsson M: **Research production in the arts and humanities. A questionnaire study of factors influencing research performance.** *Scientometrics.* 1996; **37**(3): 417–432.

[Publisher Full Text](#)

Hug SE, Ochsner M, Daniel HD: **Criteria for assessing research quality in the humanities: a Delphi study among scholars of English literature, German literature and art history.** *Res Eval.* 2013; **22**(5): 369–383.

[Publisher Full Text](#)

Hug SE, Ochsner M, Daniel HD: **A framework to Explore and Develop Criteria for Assessing Research Quality in the Humanities.** *International Journal for Education Law and Policy.* *Forthcoming.* 2014; **10**(1): 1–14.

[Reference Source](#)

Nederhof AJ, Zwaan RA, De Bruin RE, *et al.*: **Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioural sciences – a comparative study.** *Scientometrics.* 1989; **15**(5–6): 423–435.

[Publisher Full Text](#)

Ochsner M, Hug SE, Daniel HD: **Indicators for Research Quality for Evaluation of Humanities Research: Opportunities and Limitations.** *Bibliometrie - Praxis und Forschung.* 2012; **1**: 4.

[Reference Source](#)

Ochsner M, Hug SE, Daniel HD: **Four types of research in the humanities: setting the stage for research quality criteria in the humanities.** *Res Eval.* 2013; **22**(2): 79–92.

[Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:

Version 1

Reviewer Report 10 August 2015

<https://doi.org/10.5256/f1000research.6951.r9400>

© 2015 Faggiolani C et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chiara Faggiolani

Science, Technology and Society, Università degli Studi di Roma, Rome, Italy

Domenica Fioredistella Iezzi

Dipartimento di Ricerca Filosofica, Università degli Studi di Roma Tor Vergata, Rome, Italy

Consistency of title and abstract and potential interest of the subject for readers

The paper deals with a particularly relevant and controversial topic: journal ratings as predictors of articles quality in no-bibliometrics areas (architecture, arts and humanities, history and philosophy, law and sociology and political science).

The aim of this paper is to understand whether the probability of receiving positive peer reviews is influenced by having published in an independently assessed, high-ranking journal.

The title is appropriate for the content of the article and the abstract represent a suitable summary of the work. The paper is interesting and provides a useful exposure to the objectives proposed. It is well structured and clear in every step, although the article is brief and therefore omits some very important details.

Article content: Have the design, methods and analysis of the results from the study been explained and are they appropriate for the topic being studied?

The paper analyses a large dataset composed of over 11,500 research articles published in Italy in the period 2004-2010 in the areas of architecture, arts and humanities, history and philosophy, law, sociology and political sciences (non-bibliometric areas).

The authors used an ordered profit model to understand whether the probability of receiving positive peer reviews is influenced by having published in an independently assessed, high-ranking journal. The results show that independent classifications of journals may be considered as good predictors of the score assigned to individual articles.

In the manuscript, the authors use a big dataset that represents a very important first evaluation of the Italian university research. The paper is remarkable and underlines how a journal of high quality can ensure an excellent rating for a researcher. The method is suitable to achieve the goals. However, it should be highlighted that:

- 1) in the same non-bibliometric area, but in different scientific fields, monographs or articles of books rather than articles of the journal could be important, then, also kind of publication could be affect the evaluation.
- 2) the method adopted is appropriated, but it could be relevant applied clustering techniques to detect the profiles of micro-sectors (scientific fields), and find clusters of scientific fields, which although belonging to different areas have similar;
- 3) the paper deserves a greater depth to allow you to specify in detail the analysis. In fact, the authors present the results at macro level, and not enter into the details of different scientific fields. It would be interesting to have a longer version of the paper.

Conclusions: Are the conclusions sensible, balanced and justified on the basis of the results of the study?

Using a very large dataset of journal articles published in HSS, the paper proves that journal ranking may be considered as a good predictors of scores assigned to individual articles.

The probability of receiving an excellent score almost doubles when the paper is published in a top journal.

The results indicate that journal classifications may be a useful supporting tool in large evaluation exercise since it may provide reviewers with valuable information apt to support expert evaluation.

It is very important to emphasize, in agreement with *The Leiden Manifesto for research metrics* that «the best decisions are taken by combining robust statistics with sensitivity to the aim and nature of the research that is evaluated. Both quantitative and qualitative evidence are needed; each is objective in its own way. Decision-making about science must be based on high-quality processes that are informed by the highest quality data» (Hicks, *et al.*, 2015).

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 04 August 2015

<https://doi.org/10.5256/f1000research.6951.r9401>

© 2015 Zuccala A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alesia Zuccala

Royal School of Library and Information Science, University of Copenhagen, Copenhagen, Denmark

Brief Overview:

This study is situated within a national research evaluation context in Italy (i.e., designed for the selection of the best candidates for the ranks of associate and full professor) and focuses on two distinct rating exercises, one for journals and one for individual research articles published in five fields of the Social Sciences and Humanities. The authors take advantage of a fortuitous data set, and use an ordered probit-model to compare the score given by expert peer reviewers to 11,500 research articles (i.e., Excellent-A; Good-B; Fair-C; Limited-D) and the rating of the journal (classified as A or non-A) in which the individual articles were published. For all papers, a series of additional variables are taken into account, including a) language of the article, b) the scientific area of the author, c) the author's age, academic status and gender, and d) the inclusion of an international co-author. In terms of referees, an allowance is made for the possibility that he/she was international. The purpose of this controlled experiment was to test for the robustness of expert-based journal ratings by determining the probability of a paper receiving a high independent review score, where the journal in which it was published also received a high independent score.

Assessment:

I tend to agree with the first reviewer in that more background information is needed regarding the criteria behind the original two rating exercises, although I basically find that for a very short article, the statistical aspects of the methodology are quite thorough.

I have only one comment about the rating exercises. The article indicates that different panels or groups of individual experts were chosen: some were assigned to provide the journal classifications (but we do not know how many), and a separate group of others (i.e., one non-panel, plus a consensus panel) were asked to rate individual articles from the specific fields (also how many?). Neither group were said to have exchanged information; thus acted independently of one another. The peer reviewers of the individual articles were "instructed to evaluate articles only on the basis of their merit regardless of journal and of language of publication" (p. 3). Here, I am curious specifically about how much information pertaining to the journal (e.g., header; footer; abstract; citation style; volume number) had been removed prior to the article evaluation/rating procedure? Were the reviewers in this experiment truly blind to the journal's influence? In certain areas of the Humanities, where there are notably fewer A-class journals, perhaps the format of the article instantly gave away the type of journal in which it was published. This could mean that when an article had been rated as being "excellent", the reviewer was actually making a simultaneous judgement on the journal as well. I would therefore like to assume that all of the articles peer-reviewed in this experiment were non-formatted pre-prints (?); otherwise, this could be a factor which is contributing to the mutually related ratings.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 20 July 2015

<https://doi.org/10.5256/f1000research.6951.r9402>

© 2015 Williams G. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Geoffrey Williams

EvalHum Initiative and Université de Bretagne-Sud, Lorient, France

Overview

This paper sets out to exploit some of the massive amount of data collected during the Italian 2012-2014 National Scientific Habilitation (Abilitazione Scientifica Nazionale; ASN). The database created for this evaluation exercise listed the academic output of all Italian researchers applying to become associate or full professors within the Italian university of research institutes. Data was classified by output type – academic, non-academic, journal articles, books, monographs... - and also by the disciplinary section of the candidate. Interestingly, the collected data for article publications went through two independent review processes with a panel ranking journals as academic and non-academic, and giving different weighting to the academic series, and another panel carrying out peer review to look at the quality of the articles submitted.

This aim of the study described in this paper was to see if there is agreement between the rankings produced by the two panels that would allow journal rank to be a reasonable proxy for quality. It looks at a selected number of very different fields from the HSS, namely architecture, arts and humanities, history and philosophy, law and sociology and political science. The study proceeds by a statistical analysis of journals ranking and the classification of individual articles given by reviewers on a scale from A (excellent) to D (limited). It also takes into account a series of variables such as the language of production and field size as well as the discipline and age of the writer.

The results are clearly of great interest in developing new bibliometric tools for handling data in large evaluation exercises in that they show a clear correlation between journal ranking and the outcome of peer review appraisal. The outcomes indicate that those articles with the best evaluation appeared in class A journals and that thus, although peer review remains important, a journal ranking may be a good proxy for quality.

Assessment

As this is a short article, much background knowledge would be needed to fully apprehend the criteria behind the original exercise. We have little information as to either the peer review panels and the reality of the criteria they applied, or how the panel classifying journals work, and the extent that it simply followed existing classifications as that of ERIH. These are both key elements in evaluating the output as they represent sources of potential bias and the influence of normative

approaches to quality. The following comments thus concern more potential underlying biases in data collection and evaluation policy that the methodology applied here.

The first question applies to the ranking of the journals. Who are the experts and what is the danger of field bias? ERIH has come in for enormous criticism as the degree to which the experts represent a field is far from clear. In the case of this assessment, we do not know how the panels were constituted and to what extent the inclusion of a journal in an A list was free of domain bias. In France, the AERES agency had to abandon lists for the HSS as being too hotly disputed, and the field of law had simply ranked its own journals as A anyway. It is relatively easy to highlight a group of high profile, high impact journals in any field, but much more difficult to obtain clear criteria for ranking other journals as reputation measures can be fearfully biased.

The same problems arise with peer review, namely representativity of the panels and relevance of criteria to individual fields. We do not know how many reviewers read any individual paper and the extent they actually read the whole paper and were more competent to judge than the field specific peer reviewers who initially reviewed the article. The outcomes from the UK REF showed that reviewers only skimmed publications due to lack of time, given the volume of data to be treated, did these reviewers read more closely. Another issue arises from the allocation of experts, notably international experts who were chosen because the article was already deemed as having a quality potential.

Another important factor is the variety within the broad field of SSH, and even variety within disciplines. Architecture might be expected to have a more engineering dissemination profile, whereas political science and sociology can be at the end of the HSS spectrum that is closer to the sciences. The area 11, History, philosophy, pedagogy and psychology, is particularly wide as psychology can have a very different dissemination pattern to history, and is often not included or treated as borderline, within the HSS sphere. Within language, there will be great variation between areas. The oft-cited maxim that humanities researchers write books has been shown to be far too simplistic and is just an example of how broad brush strokes can hide the diversity within fields.

The article itself does point out that there can be perverse effects of an evaluation process, especially if it is normative. This may be happening here with internationalisation. Internationalisation is obviously an added value in research, but it should not be seen as a necessary prerequisite of quality. A relative lack of internationalisation is inherent in many humanities disciplines, and is particularly common in law. This is not a lack of quality, but simply due to the national orientation in the field of study. Penalising by too rigid evaluation criteria would be a bad thing as it is for evaluators to understand a field and adapt, not to try and change the field to suit their criteria.

These factors do not change the interest of the methodology adopted, but do need to be considered before making policy decisions based on outcomes. The methodology itself is thorough and opens vistas for analysing what is happening in assessment exercises and how it affects dissemination practice. I have only one minor gripe in what is otherwise a very clear and stimulating paper, and that is with data presentation. For the two datasets shown in tables 1 and 2, it would have been nice to have had averages shown so that we have an instantly visible means of comparison. This would have revealed that law beats all disciplines with a very high level of A class journals, 68% as opposed to only 39% in architecture. These differences merit comment.

The article uses sound methodology and opens perspectives for more detailed work, and hopefully a close grain analysis within disciplines that will take into account researcher motivations. This is vital as correlation at this level does not necessarily justify the use of bibliometric criteria over peer review as there are inevitably built in biases in both quantitative and qualitative approaches. Until in-depth studies of how and why researchers disseminate are carried out, the picture will always be falsified. Thus, this research opens interesting channels, and always calls for much more sociological analysis before confirming outcomes.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Comments on this article

Version 1

Reader Comment 10 Jun 2016

Alberto Baccini

The aim of (Bonaccorsi *et al.* 2015) is to “to understand whether the probability of receiving positive peer reviews is influenced by having published in an *independently assessed* high-ranking journal” [Italic added].[1] The central tenet of the paper is that there are “two independent evaluations” that can be compared.

“On the one hand, a panel of experts classified [...] academic journals in A-rated and non A-rated. The rating exercise was done on the basis of the reputation, esteem, diffusion and impact of journals, that is, on a qualitative, expert-based, reputational basis. On the other hand, we also have the rating of individual articles published in those journals, which have been done by a large number of individual referees (not panels) and summarized with a consensus agreement approach by expert panels, who however acted independently from the other panels, and without exchange of information. This peculiarity of the Italian context and the time sequence of events creates a favorable condition for carrying out a controlled experiment.”

Bonaccorsi *et al.* build a dataset by matching two different administrative databases produced, by the Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR) in their institutional activity.[2] In this comment I will argue that according to ANVUR official records,[3] the “two evaluations” cannot be considered as being “independent”, because they were reached through two intertwined procedures. As a consequence, the positive relationship among peer review evaluations and journal rating found by Bonaccorsi *et al.* cannot be considered as a sound evidence that “journal ratings are good predictors of article quality”.

1. Consider first the evaluations of individual articles. It was developed in the context of the national research assessment exercise covering the years 2004-2010 (hereinafter indicated by its

Italian acronym, VQR). The peer review process started in September 2012 and concluded in February-March 2013. It was “conducted by five Groups of Evaluation Experts (GEV in the Italian acronym), one for each [research] area”. The members of the five GEVs were nominated by the board of ANVUR,[4] an issue that will become relevant later on.

GEV members were in charge of choosing reviewers and to define evaluation criteria by considering the general ones stated by the ANVUR board (more detailed descriptions of the procedures are available in Baccini and De Nicolao 2016, Ancaiani et al. 2015). However, the first accomplished duty of all the five GEVs was the development of “rankings of journals”, by classifying journals in two or three classes. These rankings[5] were published by March 2012, before the starting of the peer review process.[6]

The peer review process was described in official documents as an “informed peer review” process,[7] because each article was evaluated by two reviewers knowing the complete metadata of the article, in particular the journal in which it was published. They knew also the rank of that journal according to the ranking developed by the GEV that called them as reviewers.[8] Bonaccorsi *et al.* wrote:

“reviewers were instructed by GEV to evaluate articles only on the basis of their merit regardless of the journal in which they are published in and of the language of publication”.

This does not correspond to the official records of the procedure. Indeed, GEV8 for Architecture “produced and used in the informed peer review the journal ranking”.[9] The GEV11 for History, philosophy, pedagogy and psychology asked reviewers to “consider also ... the quality of publication venue”. Reviewers were furnished with the “journal rankings, which, however, they were not obliged to comply”. Moreover, for articles published in journal classified as A, GEV11 might have decided also to ask only one reviewer report.[10] The GEV12 for law, explicitly stated that its journal ranking had to be considered as one source of information among the others available to reviewers for their evaluation.[11] And analogously the GEV14 for Political and social sciences considered the ranking as a way to furnish to the reviewers, “especially [to] foreign reviewers” “better information (according to the logic of informed peer review)”.[12] Only the GEV10 for antiquities, philology, literary studies, art history, established that reviewers, even though they knew the published journal ranking, had to perform a “peer review” where “the venue of publishing, the type and the language in which the research has been expressed are not factors conditioning the evaluation”.[13]

Last but not least: if two reviewers disagreed about the evaluation of an article, a consensus group decided the final evaluation. Consensus groups were generally composed by two members of the GEV, who ranked also journals. Data about the number of journal articles evaluated by consensus groups were not available (on this point see Baccini and De Nicolao 2016).

In short: we can affirm that journal rankings developed by GEVs were used by reviewers as information for evaluating articles, according to “the logic of informed peer review”. Readers have to wait a bit for understanding the relevance of this statement.. Consider now the journal ratings used by Bonaccorsi *et al.*. It was developed in the context of the National Scientific Habilitation (hereinafter ASN, according to Italian acronym). According to the Ministerial decree[14] ruling the

ASN, ANVUR was charged of producing a list of scientific journals and a list of class A journals.[15]

2. For developing the journal lists, “the working group on books and scientific journals”, coordinated by Anvur president, was constituted. It was composed by the members of the ANVUR board, and by other scholars nominated by the ANVUR board and organized in five different Panels,[16] defined in reference to the research areas already described for the GEVs . Each panel was composed by 4 members, with the only exception of Panel10 composed by 8 members.[17] Recall that ANVUR board had already nominated the GEVs. Works and criteria adopted by the Panels for the ratings were only summarily described in short final reports, with the exception of Panel10 that never published a report for its work.[18]

The ministerial decree stated that for developing the journal lists, ANVUR “avails upon the GEVs for the VQR”. [19] According to this provision, ANVUR stated that it “retains to get opinions from GEV”[20] and that it will deliver the final lists of journals after having considered “the observations and proposals of the GEVs”. [21]

5th October 2012, some weeks after the definitive publication of journal lists, and for replying to a growing mass of criticisms (Mazzotti 2012, Baccini 2016), ANVUR published a document where it was clearly stated that Class A lists developed by the Panels, “as suggested by the Ministerial Decree 76/2012, were sent to the GEVs by asking their opinions”. The only exception was for GEV10. The opinions expressed by the GEVs were positive, with “small modifications” required by GEV11 and by GEV14.[22]

According to available official records, GEVs not only gave their final opinion, but interacted in many ways with panels.[23] These are the interactions documented in the publicly available records:

- all the five Panels received the appropriate journal rankings developed by the GEVs for the VQR, and used it as the starting source of information for their work.[24] Panels received the mandatory indication of considering also another source of information: the ratings possibly developed by academic societies.[25] ANVUR President sent a mail to academic societies by soliciting the production of these journal ratings, and by suggesting that to this end they “consider the work of classification developed by the GEVs”;[26]
- Panel8 (architecture) had at least one meeting with the coordinator of the GEV8 who “illustrated criteria and goals” of the VQR ranking. Indeed, Panel8 adopted the Class A list developed by GEV8 and added to this list a “limited number” of other journals;[27]

Panel14 (political and social sciences) explicitly indicated that the GEV14 coordinator and one of the member of ANVUR board participated to the meetings. The GEV14 coordinator was explicitly charged to act as “trait d’union” between Panel14 and GEV14.[28] As for the preceding GEV, the Class A list developed by the GEV14 was confirmed by Panel14 with some new insertions.[29]

3. Consider now the “time sequence of events”. The journal rankings for the VQR developed by the GEVs were ready by March 2012; the journal ratings developed by Panels for the ASN were ready by July-August 2012; the informed peer review began in late September 2012. Reviewers called to

evaluate papers for the VQR, at least Italian reviewers, knew that journal ratings were developed for the ASN, and that they were only partially different from the ones developed by the GEVs for the VQR. Reviewers, "in the logic of informed peer review" might have used the ratings for the ASN as information for their evaluations of articles. This short circuit was testified by GEV10 by explicitly writing in its final report that the publication of the class A list was an "element of disruption" for the VQR. [30]

This is another reasons for which the two evaluations used by the Bonaccorsi et al. cannot be considered as "independent".

4. This comment documents that according to available official records, data used by Bonaccorsi et al. about journal ratings and scores of individual articles cannot be considered as generated by "two independent evaluations", whatever the meaning of the adjective "independent". Indeed the two evaluations were intertwined in various ways. They were organized by a same group of scholars: the ANVUR board that was in charge also of choosing the members of both the GEVs and the Panels. ANVUR board, GEVs and Panels interacted during the development of ASN journal ratings; these ratings were developed starting by the VQR journal rankings developed by GEVs. Finally reviewers, in charge of evaluating articles for the VQR, were asked to consider the venues of publication and in particular the journal rankings. It is therefore hardly surprising that Bonaccorsi et al. found that "the probability [of articles] of receiving a better scores grows with the quality profile of the journal". Bonaccorsi et al.'s paper did not represent a sound addition to the body of literature on journal classification.

References

- Ancaiani, Alessio, Alberto F. Anfossi, Anna Barbara, Sergio Benedetto, Brigida Blasi, Valentina Carletti, Tindaro Cicero, Alberto Cioffi, Filippo Costa, Giovanna Colizza, Marco Costantini, Fabio di Cristina, Antonio Ferrara, Rosa M. Lacatena, Marco Malgarini, Irene Mazzotta, Carmela A. Nappi, Sandra Romagnosi, and Serena Sileoni. 2015. "Evaluating scientific research in Italy: The 2004–10 research evaluation exercise." *Research Evaluation* no. 24 (3):242-255. doi: 10.1093/reseval/rvv008.
- Baccini, A. 2016. "Napoléon et l'évaluation bibliométrique de la recherche. Considérations sur la réforme de l'université et sur l'action de l'agence national d'évaluation en Italie." *Canadian Journal of Information and Library Science-Revue Canadienne des Sciences de l'Information et de Bibliothéconomie* no. 40 (1):37-57. doi: 10.1353/ils.2016.0003.
- Baccini, Alberto, and Giuseppe De Nicolao. 2016. "Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise." *Scientometrics*:1-21. doi: 10.1007/s11192-016-1929-y.
- Bonaccorsi, A, T Cicero, A Ferrara, and M Malgarini. 2015. *Journal ratings as predictors of articles quality in Arts, Humanities and Social Sciences: an analysis based on the Italian Research Evaluation Exercise [version 1; referees: 3 approved]*. Vol. 4.
- Mazzotti, Massimo. 2012. Listing wildly. *Times Higher Education*, 08/11.

NOTES

[1] Raw data were not disclosed to other scholars, not even in a feasible anonymized form. The link provided in the section “Data availability” of the paper does not contain any more data concerning the article. I will provide below the direct links to the relevant documents contained in the ANVUR website. In many cases I found the links by using Wayback machine <https://archive.org/web/>.

[2] Bonaccorsi et al. inappropriately described that research activity as a “controlled experiment”.

[3] In this comment I will provide my own English translations for quotations.

[4] Andrea Bonaccorsi was one of the seven members of the board of ANVUR.

[5] Hereinafter the words “ranking/rankings” refers only to the journal ranking developed by GEVs for VQR, and the words “rating/ratings” refers only to the journal classification developed by Panels for National Scientific Habilitation and used by Bonaccorsi et al..

[6]

<http://www.anvur.org/rapporto/main.php?paragraph=3.1.1&cap=My4xLjEuIFtlbV1MYSBjbGFzc2lmaWNhemlrbmUg>

[7] http://www.anvur.org/rapporto/files/VQR2004-2010_RapportoFinale_parteprima.pdf, p. 5 and *passim*.

[8] http://www.anvur.org/rapporto/files/VQR2004-2010_RapportoFinale_parteprima.pdf, p. 21.

[9] *ibidem*

[10] http://www.anvur.org/rapporto/files/Area11/VQR2004-2010_Area11_RapportoFinale.pdf, pp. 33-34.

[11] http://www.anvur.org/rapporto/files/Area12/VQR2004-2010_Area12_RapportoFinale.pdf, pp. 32-34. And also http://www.anvur.org/rapporto/files/Area12/VQR2004-2010_Area12_Appendici.pdf

[12] http://www.anvur.org/rapporto/files/Area14/VQR2004-2010_Area14_RapportoFinale.pdf, p. 17.

[13] http://www.anvur.org/rapporto/files/Area10/VQR2004-2010_Area10_Appendici.pdf. Appendice B, p. 7. The effectiveness of this recommendation cannot be verified.

[14] Ministerial decree n. 76/2012, art. 6.6. <http://attiministeriali.miur.it/anno-2012/giugno/dm-07062012.aspx>

[15] Annex b of the Ministerial Decree, http://attiministeriali.miur.it/media/192907/dm_07_06_2012_allegatob.pdf. ANVUR stated its duties in the deliberation n. 50/2012; http://www.anvur.org/attachments/article/252/Delibera50_12.pdf

[16] This provision permitted to the members of the ANVUR board to participate to the meetings of the panels. Since the minutes of the meetings were not published, it is impossible to know if ANVUR board members really participated to the meetings and how they interacted with the other panelists in the final decisions.

[17] The composition of the panels was defined in these ANVUR deliberations:

http://www.anvur.org/attachments/article/422/delibera55_12.pdf;

http://www.anvur.org/attachments/article/422/delibera58_12.pdf;

http://www.anvur.org/attachments/article/422/delibera63_12.pdf.

[18] http://www.anvur.org/attachments/article/254/Relazionefinale_GdLArea08.pdf ;

http://www.anvur.org/attachments/article/254/Relazionefinale_GdLArea11.pdf;

http://www.anvur.org/attachments/article/254/relazionefinale_gdlarea12.pdf;

http://www.anvur.org/attachments/article/254/relazionefinale_gdlarea14.pdf.

[19] <http://attiministeriali.miur.it/anno-2012/giugno/dm-07062012.aspx>, Annex B, section 2.

[20] http://www.anvur.org/attachments/article/252/Delibera50_12.pdf , Art. 11.6.

[21] *Ibidem*, art 12.3.

[22] GEV14's request of modification was limited to the sub-field of political sciences

http://www.roars.it/online/wp-content/uploads/2012/10/chiarimenti_riviste_scientifiche.pdf, pp. 1-

2. On 24th September 2012, ANVUR published a document, where it was clearly stated that "the lists were previously submitted to the opinion of the GEV". After 4 days the document was replaced by another one. The only difference between the two documents was that the sentence quoted above was dropped in the second document. Copy of the documents and a discussion is available here: <http://www.roars.it/online/lenigmistica-di-anvur-trovate-le-differenze/>

[23] In an ANVUR document of January 2013, it was clearly stated that ANVUR developed the journal lists "by employing ... the GEVs and the working group [on books and scientific journals]. The document was an undated pdf. The properties of the document registered it as last modified on 13th January 2013, and as authored by "Bonaccorsi"

<http://www.anvur.org/attachments/article/252/riviste.pdf>, p. 2.

[24] http://www.roars.it/online/wp-content/uploads/2012/09/chiarimenti_riviste_classea_0.pdf

[25] Art. 12.4 of http://www.anvur.org/attachments/article/252/Delibera50_12.pdf.

[26] For an example of such a letter: <http://www.glottologia.org/wp-content/uploads/2012/07/Documento-10-1-ANVUR.pdf>.

[27] The complete description is available here:

<http://www.anvur.org/attachments/article/252/riviste.pdf>

[28] http://www.anvur.org/attachments/article/254/relazionefinale_gdlarea14.pdf, p. 1

[29] The complete description is available here:
<http://www.anvur.org/attachments/article/252/riviste.pdf>.

[30] http://www.anvur.org/rapporto/files/Area10/VQR2004-2010_Area10_RapportoFinale.pdf, p. 26.

Competing Interests: None.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research