

Issue 15 . No.2 2015

# Bibliometric Analysis Of Publications Related To Innovation Projects And Their Management In Scopus, In The Period 2001-2011

Juan C. Hernandez, School of Pharmacy, University of Costa Rica, Costa Rica

### ABSTRACT

Introduction: innovation projects and their management as a tool to contribute to the development and competitiveness of organizations, regions and countries are becoming increasingly relevant, mainly for countries with limited resources. Promoting its development and application requires delving deeper into the behavior of scientific production on the subject worldwide. **Objectives:** to know the current state of the management of innovation projects from the characterization of the bibliographic records of scientific articles related to the subject, indexed in the Scopus database and published in the period 2001-2011.

Methods: a bibliometric analysis of publications on innovation projects and their management indexed by the Scopus database was carried out, corresponding to the period 2001-2011. A total of 720 articles were identified; Productivity was studied by years and countries, the most productive authors and institutions, and the levels of collaboration between authors. Results: a trend was observed to increase the number of publications led by developed countries, and a low representation of Latin American countries.



**Keywords:** innovation projects, innovation project management, bibliometrics, bibliometric indicators, scientific production.

### **INTRODUCTION**

Innovation is perceived as a source of development and competitiveness of nations, regions and organizations, <sup>1</sup> which is why it has gained increasing importance in theoretical models of economic growth and in business literature.

The study of innovation as a process, as well as the tools to manage it and the indicators for its evaluation, are the subject of countless research carried out from different perspectives and at different levels of socioeconomic development.<sup>23</sup>

An efficient tool to manage innovation is project management, which has initially been practiced since ancient times, and has been fundamentally related to engineering projects for the construction of civil works. <sup>4</sup>

At the end of the 19th century and the first half of the 20th century, the gradual development of administration and planning techniques represented by *Frederick Taylor* (1856-1915) and *Henry Gantt* (1861-1919), the emerging development of technology, together with the development of technical to increase the effectiveness of military projects in the 1950s, among which the Critical Path Method (*CPM*) developed by *Dupont Corporation* in 1957 stands out, as well as the program evaluation and review technique (*Program Evaluation and Review Technique or PERT*), *used for the Polaris* Project developed by the United States Army in 1958, and which spread rapidly to other types of industry, <sup>5</sup> led to project management becoming more widespread after the Second World War. became a research discipline. <sup>6</sup>



In the business field, the concept of project management in this sense was coined around the middle of the 20th century in the United States under the term known as "*Project Management*", translated into Spanish by *Heredia*<sup>4</sup> as Integrated Project Management (DIP). defined in the *Project Management Body of Knowledge*, of the *Project Management Institute* of the United States as "the art of directing and coordinating human and material resources, throughout the life cycle of a project, through the use of current m techniques. *anagement*, to achieve the preset objectives of scope, cost, deadline, quality and satisfaction of the participants or parties interested in the project.<sup>7</sup> In Abstract, it is about systematizing the management techniques and forms of organization appropriate to face complex operations that are very difficult to master by applying classic management systems.

In the last three decades, project management has evolved significantly, and within it the management of innovation projects, fundamentally motivated by the demands of the accelerated changes in all global socioeconomic processes of recent years, which have made innovation a driving force of development in all sectors and spheres of society.

In relation to the management of innovation projects, although there is no consensus regarding a conceptualization from a theoretical point of view, different considerations have been made that are very useful from a methodological point of view, among which are the proposal of *Ramírez* (2006), for whom the management of innovation projects is "the set of technical and administrative acts that allow an innovative idea to be converted into a viable, sustainable project that is susceptible to equitable *ex post* evaluation ." <sup>8</sup>



In the Innovation Manual: Practical Guide to R&D&I Management of SMEs (2007), project management seeks to apply the knowledge, skills, tools and techniques available in SMEs, so that they can direct and coordinate ongoing operations, to achieve compliance in time and cost with the set objectives. <sup>9</sup>

The theoretical and methodological support for the management of innovation projects has been developed fundamentally in developed countries in North America, Europe and Asia, which have created various tools to achieve effective management of innovation projects, fundamentally on a business scale and with emphasis on technological innovation.

However, the expanded conception of innovation that also extends to organizational <sup>10</sup> and social <sup>11</sup> innovation and exceeds the limits of the organization, leads to the continuous evolution of this discipline, which although it has been promoted from developed countries, is vital that it be promoted in countries with low levels of socioeconomic development.

To contribute to the development of the management of innovation projects, it is necessary to deepen the scientific knowledge generated around this topic, which implies knowing the status achieved on the topic reflected in the literature generated, a process that requires the participation of other disciplines. scientific such as Bibliometrics.

Bibliometrics is the science that aims to study quantitative data from scientific publications. <sup>12-15</sup> According to *Spinak*, it studies the organization of scientific and technological sectors based on bibliographic sources to identify the authors, their relationships, and their trends. <sup>16</sup> According to *Pérez Matos*, bibliometrics is a tool capable of determining phenomena, trends and regularities that occur in the



scientific field based on its literature, regardless of the fact that many knowledge and elements of scientific phenomena are not written down. <sup>17</sup> In such a way that bibliometric studies acquire increasing relevance for the scientific community due to their valuable contributions to the knowledge of the state of an area or a research topic. <sup>13.17</sup>

This work constitutes a bibliometric study of global scientific production that addresses innovation projects and their management. Its objective is to know the current state of the management of innovation projects from the characterization of the bibliographic records of scientific articles related to the subject, indexed in the Scopus database and published in the period 2001-2011, determining from of bibliometric indicators <sup>18-20</sup> the behavior of publications by year, the authors, institutions and countries with greater productivity, the cooperative relationships established, as well as the journals that publish the most on the subject.

#### **METHODS**

The Scopus database published by Elsevier was used as a primary source of information. It is a bibliographic database of summaries and citations of scientific journal articles. It covers approximately 19,000 titles from more than 5,000 international publishers, including 16,500 peer-reviewed journals in science, technology, medicine, social sciences, arts and humanities. Allows searching of scientific web pages using Scirus, Elsevier and patent databases.

Elsevier's Scopus database was accessed through the open access SCImago Country and Journal Rank (SJR) portal. The search was carried out in the period between 2001 to 2011, and initially the terms innovation projects and innovation



project management were used for the search strategy, with which only one record was obtained, so it was decided to use their English equivalents *innovation Project*, for which 720 records were retrieved, and *innovation project management*, with 159 records. Subsequently, possible overlaps in the identified records were checked and it was found that the 159 records that responded to *innovation project management* were contained in the total of records identified for the terms *innovation project*, leaving a total of 720 for the study.

To compile and process the data, Microsoft Excel software was used (a spreadsheet or electronic tabulator type program that allows operations with numbers organized in a grid, from simple operations to complex statistical calculations, and the preparation of tables and graphs. ) and ToolInf (analysis tool developed by the BioMundi Consultancy of Cuba, which allows the homogenization and counting of data and preparation of matrices). The information for the study of the journals on the selected topics was structured in the following fields: authors, journals, country, topic, year and institution.

Finally, the files obtained were taken to Ucinet and within this NetDraw (<u>http://www.analytictech.com/ucinet/trial.htm</u>) was used to obtain co-occurrence matrices between two variables, which It allowed mapping, editing and analyzing social matrices and visualizing them.

Next, the indicators that will be used in the study were operationally defined; among them: productivity by years (total number of articles published for each year included in the study), authorial productivity (number of articles signed by author), co-authorship (works that are produced by two or more authors in institutions inside and outside the country), productivity by institutions (total of



articles produced by the institution to which the first author belongs in the period of time studied) and productivity by countries (total of articles produced in institutions based in the country in the period studied). Finally, all the analysis of the results obtained was carried out. In addition, the method of classic documentary analysis was used with the objective of detecting and analyzing the sources of information to obtain the methodological theoretical references of the topic, that is, the state of the art of innovation projects and their management, as well as background of bibliometric studies carried out on the topic.

For the study, the data included in the references were considered valid. The veracity of bibliographic citations was only verified in those in which the available information was insufficient and it was necessary to complete some information. If any erroneous data was found during this process, it was corrected. There was little general coverage, over the years, of Latin American publications on the subject in the database under analysis.

#### RESULTS

A bibliographic review was carried out on the topic investigated, with the purpose of identifying antecedents of bibliometric studies carried out in the field of innovation projects and their management. As a result, a growing development of theoretical research in this matter could be observed; However, no bibliometric study was found referring to this topic that analyzed the behavior of scientific production in this area, which substantiates the need to carry out this type of study. In the bibliometric study of the topic, a total of 720 articles published in the Scopus database in the period from 2001 to 2011 were identified; of them, 159 dedicated



specifically to the management of innovation projects. Its distribution by year allowed us to evaluate the trends in research on this topic (Fig. 1), and a gradual increase in publications was observed, in which the highest productivity was obtained in 2007, which is an indication of the development gradually achieved this matter. Taking into account the coefficient of determination ( $\mathbb{R}^2$ ) of the second-order polynomial trend line, a tendency to increase in the number of articles related to these topics to be published in the next five years is reflected.

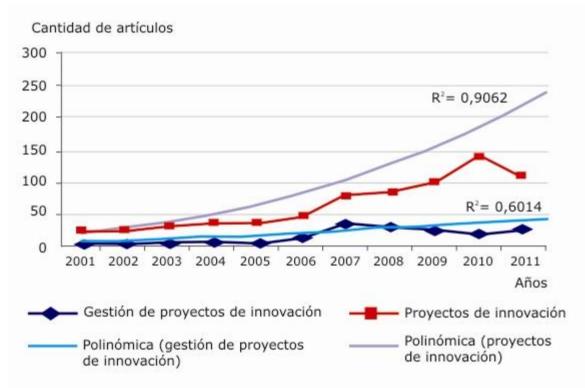


Fig. 1. Tendencia de las publicaciones de la gestión de proyectos de innovación en la base de datos Scopus.

#### PRODUCTIVITY BY AUTHORS



To analyze the productivity of the authors, the signatures of the works were initially used as a quantifiable indicator, which is commonly used to establish the most active research core and identify the most productive researchers and their professional origin. For this, the authors were listed in descending order according to the number of articles produced, and a total of 651 were identified. As this indicator (according to *Lotka*, 1926; *Price*, 1963, cited by *Pacheco* and *Milanés*, 2009) as It would only serve to get a first approximation to the topic, since this variable does not seem to fit a linear additive model, but rather a multiplicative one, it was recommended to use another criterion that allows better discrimination between the total number of identified authors, so to To select the most productive authors, it was determined to apply Price's Square Root Law, which establishes that the square root of the total number of authors produces 50% of what is written, and the remaining 50% is produced by all the others. authors. <sup>21</sup> By applying the square root to 651, a nucleus of 25 authors was obtained, which are listed in Table <u>1</u>.



#### Tabla 1. Productividad autoral en el tema proyectos de innovación

No.	Autores	No. artículos		
1	Chen Y.	4		
2	Keizer J. A.	4		
3	Lettl C.	4		
4	Chen J.	3		
5	Inoue M.	3		
6	Rotshtein A. P.	3		
7	Stevens E.	3		
8	Swan J.	3		
9	Van Riel A. C. R.	3		
10	Wang Y.	3		
11	Zhang C.	3		
12	Blindenbach-Driessen F.	2		
13	Burdick W. P.	2		
14	Bygstad B.	2		
15	Christiansen J. K.	2		
16	Ciabuschi F.	2		
17	Dayan M.	2		
18	Golovatchev J.	2		
19	Hoegl M.	2		
20	Hommels A.	2		
21	HU W.	2		
22	Huang Y. S.	2		
23	Hubbard S. M.	2		
24	Kapsali M.	2		
25	Lechler T.	2		

Regarding the most productive authors, those with the greatest number of articles as main authors were investigated in depth, among which the following stood out:

• *Junjie Chen*: from the College of Computing and Software, Taiyuan University of Technology, China. He is the most productive author on the topics of



innovation project management and innovation projects. The subject areas where he publishes are biochemistry, genetics and molecular biology; computer's science; medicine; business, administration and accounting; engineering; science decision; agriculture and life sciences; energy, environment, among others, with an *h index* of 28.

Jimme A. Keizer and Christopher Lettl : they are the most productive authors on the topic of innovation projects. The first is from the Faculty of Technology Management at the Technical University of Eindhoven, in Norway. The subject where publishes business. administration areas he are and accounting; engineering; social Sciences; psychology and computer science, with an h-index of 5. The second is from the Institute for Entrepreneurship and Innovation, at the Vienna Economic University, in Austria. The subject areas administration where he publishes business. are and accounting; engineering; science decision; economics, econometrics and finance; biochemistry, genetics and molecular biology; psychology and agriculture, and biological sciences, with an *h* index of 7.

#### PRODUCTIVITY BY COUNTRY

Among the most prominent countries in scientific production on the subject with more than 100 documents each during the period studied were the United States and China. In a second block with more than 50 articles were the United Kingdom, Germany and Holland. 50% of the publications analyzed were concentrated among these five countries (Fig. 2). In contrast to the high levels of scientific production



related to this topic achieved by developed countries, the low levels achieved by Latin American countries were observed, with only 17 articles, which represented 0.02% of the total publications, and the most A significant contribution was made by Brazil.

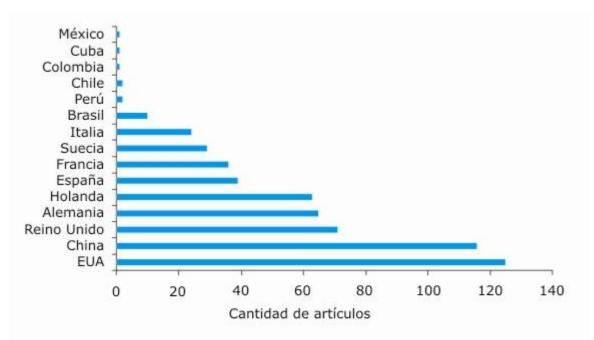


Fig. 2. Distribución de artículos sobre los proyectos de innovación y su gestión en diferentes países.

#### PRODUCTIVITY BY INSTITUTIONS

Figure  $\underline{3}$  illustrates the distribution by countries with the greatest number of institutions that have published on the subject in the indicated period, where the most representative were China, Holland and the United States, followed by the United Kingdom, Germany and Spain, a result that is corresponds to the productivity indicator by country previously analyzed.



The analysis of the most productive institutions showed that 70.5% of the publications referring to the topic come from universities. The most productive were *Maastricht University* and *Wageninger University* in the Netherlands; the *Rensselaer Polytechnic Institute*, in the USA; Northwestern *Polytechnical University*, China, and *Aalborg University Copenhagen*, Denmark, represented by thicker lines in Figure 3.

### AUTHORAL COLLABORATION

The records identified in the present study showed that 82% of the publications referring to the topic were produced by a maximum of three authors, which could be related to the fact that most of the research in this field responded to theoretical works, which —as *J. Sylvan Kats* demonstrated— produced articles with few authors compared to experimental works.<sup>22</sup>

To delve deeper into the behavior of author collaboration, the analysis of the coauthorship network was carried out, based on the visual representation of those authors who collaborated on two or more research works, which allowed the identification of 28 fundamental components (Fig. 4). 85% of the components were made up of two (17; 60%) and three (7; 25%) authors. The main components were composed of five authors; The first was made up of: *J. Inoue*, *H. Murakami*, *K. Mahamud*, *G. Wu* and *M. Hasegawa*, all from the *Communications Research Laboratory* of Japan, and the second was a reflection of the collaboration between several universities, made up of: *J. Swan*, *S. Newell* and *A. Goussevskaia*, *University of Warwick*; *M. Robertson*, of the *University of* 



*London*, and *M. Bresnen*, of the *University of Leicester*, all from the United Kingdom. The components that reflected the co-authorship relationships of the three most productive authors analyzed previously are differentiated with the color green; Of them, the Austrian *C. Lettl* was the one who exhibited the greatest number of relationships.

In general, the collaborative relationships that were established in this field were led, although not in all cases, by authors who were in the group with the highest authorial productivity, with a predominance of collaboration between authors from national institutions.

### MAGAZINES THAT PUBLISH THE MOST ON INNOVATION PROJECTS AND THEIR MANAGEMENT

The journals that publish the most on the topic were selected, and the *ranking* of all was established with the support of the *SCImago Journal Rank* (SJR) and H indices  $\cdot$  <sup>19</sup> Of the total of journals, nine are located in the first quartile, ordered by the SJR indicator, where the *Journal of Product Innovation Management* and the International *Journal of Project Management* are the ones that published the greatest number of articles related to the topic in the period analyzed, which is why they were placed first and third, respectively, in the *ranking* of journals established for the present study (<u>table 2</u>)



Publicación	Cantidad de artículos	SRJ	H index
Journal of Product Innovation Management	13	1,93	66
Technological Forecasting and Social Change	6	0,97	41
International Journal of Project Management	18	0,84	52
Construction Management and Economics	7	0,64	38
European Journal of Innovation Management	6	0,63	15
International Journal of Innovation Management	5	0,56	9
International Journal of Technology Management	8	0,37	29
International Journal of Technology Intelligence and Planning	6	0,19	3
Advanced Materials Research	8	0,14	10

 
 Tabla 2. Revistas científicas que más publican en el tema en la base de datos Scopus

Fuente: SCImago Journal & Country Rank a partir de datos Scopus.

### MAIN CONTRIBUTIONS OF THE RESULTS OF THE ANALYSIS AND FUTURE LINES OF RESEARCH

The use of bibliometric indicators for the evaluation of science allows for an indepth analysis from the quantitative and qualitative point of view of the behavior of science through its scientific production.

The evaluation of the behavior of the subject related to innovation and management projects through a bibliometric study of the publications in the



Scopus database in the period 2001-2011, has shown the evolution achieved by this discipline to date and the trend to continue developing in the coming years; It also provides a working tool for researchers in this field taking into account that no references of similar studies applied to the topic were found.

The study confirms that global leadership in this area is held by developed countries led by the United States, China and the United Kingdom, followed by other European and Asian countries, while poor and developing countries have a low level of expressed representation. through its limited scientific production, which could be related to the possibility that some of the research and knowledge production activities in most developing countries appear in "gray literature publications". We must also recognize a language bias with a strong influence of English and accept that not all scientific production has the same opportunities for publication and not all publications have the same possibilities of integrating existing databases, among other elements that affect the low scientific production underdeveloped countries, which corroborates what was stated of by OECD/ECLAC (2011) in relation to the increase in the technological gap between developed countries and those still developing is fundamentally linked to the permanence of the concentration of the generation and absorption of knowledge essentially in developed countries. <sup>23</sup>

The identification of the core of authors with greater productivity, the institutions in which they are located, within which universities constitute the center of the development of research in this field, the main journals that publish on the topic studied, as well as the behavior of Authorial collaboration, which shows a tendency towards cooperation between authors from national institutions with a



predominance of intra-institutional collaboration, constitute basic elements to guide the course of research in this area.

Taking into account that this study constitutes a first approach to bibliometric studies related to the topic, as well as the limitations identified in its implementation, it is recommended as future lines of research to develop bibliometric studies of the topic covering a greater interval of time and bibliometric indicators. , and carry out bibliometric studies on this topic in other databases and directories such as Scielo, Redalyc, Latindex and others.

### REFERENCES

1. Apaza C. Comparative study on reform processes of public structures in selected countries. Support for the structural innovation project of the executive branch. Organization of American States of the Republic of Paraguay [Internet]. 2011 [cited July 2, 2012]. Available at : http://www.oas.org/es/sap/docs/dgpe/EstudioComparativo2011.pdf

2. Lundvall B, Vang J, Joseph Y, Chaminade C. Bridging Innovation System Research and Development Studies: challenges and research opportunities. Senegal: 7th Globelics Conference; 2009.

3. Loet L. Indicators of Innovation in a Knowledge-based Economy. Cybermetrics[Internet]. 2001[cited September 5, 2011]:5(1). Availableat: <a href="http://www.cindoc.csic.es/cybermetrics/articulos/v5i1p2.html">http://www.cindoc.csic.es/cybermetrics/articulos/v5i1p2.html</a>

4. University Group of Integrated Project Management. Introduction to Integrated Project Management. Notes booklet. ISPJAE. Havana: Working document. 2013.



5. Vargas C. TimeLine of project management. Business Management Bulletin [Internet]. 2012 [cited June 16, 2012]. Available at: <u>http://bge.zoomblog.com</u>

6. Bosch H. Technology Management. Reading room Science, Technology, Society and Innovation. OEI; 2000.

Project Management Institute (PMI). Guide to project management fundamentals. PMI Publications [Internet]. USA: 2008 [cited June 16, 2012]. Available at <u>: http://www.pmi.org</u>

8. Ramírez G. Innovation project management. In: Innovation management: an updated vision for the Ibero-American context. Academia Publishing. Havana: 2006.

9. European Business and Innovation Center. Innovation manual: practical guide to R&D&I management for SMEs [Internet]. Madrid: 2007 [cited September 5, 2011];43:224-40. Available

at: <u>http://www.camaracr.org/internacionalizar/informacion-internacional/detalle-de</u> <u>-herramienta/cc/manual-de-innovacion-guia-practica-de-gestion-de-la-idi-para-</u> pymes /ccac/show/Content/

10. Faloh R. Organizational Innovation. In: Innovation management: an updated vision for the Ibero-American context. Havana: Editorial Academia; 2006;5:59-76. 11. Echeverría J, Merino L. Paradigm change in innovation studies: the social turn of European innovation policies. ARBOR Science, Thought and Culture [Internet]. 2011 [cited July 2, 2012]. Available at: http://arbor.revistas.csic.es/index.php/arbor/article/viewArticle/1387

12. Spinak E. Quantitative analyzes of scientific literature and their validity for judging Latin American Production. Bull PAHO. 1995;29(4):352-9.



13. Camps D. Bibliometric study of case material articles published in the Spanish Journal of Pathology, 2005-2009 [Internet]. Rev Esp Patol. 2010 [cited February 7, 2013];43(4):196-200. Available at: <u>http://www.elsevier</u>

14. Villar F, Estrada JM, Pérez C, Rebollo MJ. Bibliometric study of the original articles of the Spanish Journal of Public Health (1991-2000). Third part: Analysis of **REFERENCES**. Rev Esp Sal Públ. 2007;81(3):247-59.

15. Pacheco J, Milanés Y. Evaluation of science and bibliometric studies [Internet]. Rev electr SIRIVS. 2009 [cited June 8, 2012] Available at: <u>http://www.unmsm.edu.pe/veterinaria/files/evaluacion\_de\_la\_ciencia.pdf</u>

16. Spinak E. Scientometric indicators [Internet]. Acimed. 2001 [cited February13, 2013];9(Suppl.):42-9. Available

at: http://bvs.sld.cu/revistas/aci/vol9\_s\_01/sci0 7200.htm

17. Pérez N. Bibliography, bibliometrics and related sciences [Internet]. ACIMED. 2002 [cited June 8, 2012];10(3). Available at: <u>http://bvs.sld.cu/revistas/aci/vol10\_3\_02/Aci012002.htm</u>

18. Gómez-García A, Ramiro MT, Ariza T, Granados MR. Bibliometric study of Education XXI. Education XXI [Internet]. 2012 [cited February 7, 2013];15(1):17-41. Available at: http://redalyc.uaemex.mx/redalyc/pdf/120/12016344016.pdf

19. Pratt AM. Training module for the collection and analysis of product indicators of science and technology activities. Inter-American Development Bank [Internet]. 2009 [cited June 9, 2012];12-34. Available at: <u>http://12docs.politicascti.net/.../Doc%2007%20-%20capacitacion%20prat.pdf</u>
20. González Guitián MV, by Zayas Pérez MR. Knowledge audits. Domain analysis in Scopus and WoK databases. Rev Interam Librarian. 2012;35(1):17-25.



ISSN: 1533-1520Eco-Fin, Accounting, and Statistics21. Morejón Bravo Y. Scientific production on demography: Its behavior in<br/>Scopus 1999-2009. Havana: 10th National Workshop for Updating and Exchange<br/>of Experiences in Sciences, Technologies, Information Management and<br/>Knowledge Management of Scientific Poles . 2011.

22. Arencibia Jorge R, Moya Anegón F. The evaluation of scientific research: a theoretical approach from Scientometrics [Internet]. ACIMED. 2008 [cited March 2013];17(4). Available

at: <u>http://scielo.sld.cu/scielo.php?script=sci\_arttext&pid=S1024-</u> 94352008000400004



# Scales And Lists For Evaluating The Quality Of Scientific Studies

Mei L. Chen, College of Engineering, National Taiwan University, Taiwan

### ABSTRACT

The objective of this study was to identify scales for evaluating the methodological quality of scientific articles and checklists for the quality of information in research in the health area. Two basic procedures were developed: a) systematic review of the scientific literature in the databases Web of Science, Journals@ovid, Science Direct, Scopus, SportDiscus, Mary Ann Liebert and Oxford Journals Online, with a selection of articles published in recent years. five years and indexed in the English language; b) bibliometric review in the references of the articles selected in the systematic review, without definition of time or language. Different studies were selected that represented 14 scales and their modifications, and also 11 lists used. It can be concluded that the scales and lists differ from each other in relation to the number of *items*, validity, reliability and scoring margins. Most of the scales and *items* present psychometric properties of validity and reliability, and are applicable to revisionary studies, mainly meta-analytic, both in the search for methodological quality and the quality of information.

Keywords : scales, health research evaluation, methodology, information.

### INTRODUCTION

The evaluation of the quality of scientific studies can be considered essential in the production and selection process of scientific literature in health. According



to *Verhagen*, <sup>1</sup> the evaluation of methodological quality considers internal validity, which refers to the analysis of the ability to adequately measure what was proposed, and external validity, which refers to the analysis of statistical hypotheses and the generalization of the results. outcomes for the population of interest; Furthermore, it allows us to analyze the transparency in the description of the objectives, the importance of the sample size to detect the clinical effect investigated and the presentation of the results. <sup>2,3</sup> This evaluation can be carried out by checklists and evaluation scales. Checklists are useful when they provide guidance and information that should be included in the reports of randomized clinical trials (RCTs). Evaluation scales provide a quantitative index of the methodological quality of RCTs <sup>4</sup> and have the advantage that they can be easily replicated and formally incorporated in the review, through peers and in systematic comments; but it also has disadvantages such as the lack of tests that decide the inclusion and exclusion of the *items* and their numerical scores attached to each of the evaluated elements.

The scales and checklists include *items* that measure the quality of the studies. The scales provide responses for the individual *items* that are processed and can offer global scores that provide points that classify the methodological quality, such as the Jadad scale, where this is the most used in RCTs, which provides a score that classifies the study as weak (0 points) to good (5 points); or the PEDro scale, which is applied in experimental studies and scores according to the presence of indicators of the quality of the evidence presented (1 point) or the absence of these indicators (0 points), up to a total score of 10 points. It should be noted that a single quality score suggests facilitation in interpretation, while some guidelines



must be followed for the evaluation of adequate performance in psychometric tests of reliability, content validity, construct validity and concurrent validity, among others, which are considered essential in the qualification process of current scientific literature. <sup>23</sup>

In a study carried out in the 1990s, 25 scales for evaluating the quality of primary studies were identified, but only one scale had been developed following consolidated methodological procedures. <sup>4</sup> Currently, various scales and lists have been produced with the purpose of increasing the methodological and information quality of different types of research in the health area, such as non-randomized clinical studies, observation and systematic review. This increase has provided important tools for researchers, scientific editors and readers of different scientific areas, since the identification of valid and reliable scales on a specific topic can minimize the chances of errors in determining the quality of scientific literature, <sup>5</sup> in executing a study and verifying the application of the results. <sup>3.6</sup>

However, many difficulties inherent to the problems in research in the area of health are still revealed in the process of evaluating the quality of the studies, such as, for example, the difficulties that partners of scientific newspapers encounter on a daily basis. inconsistencies between objectives, procedures and results, which in part have their origin in inconsistent or inadequately selected theoretical bases; or the difficulties encountered by ethical analysts who, despite the apparent care for human dignity proposed by the researchers, end up confronting the economic concerns of the study's funders, such as, for example, the definition of the sample, the indiscriminate use of the placebo or of the *wash-out*, the weaknesses in the



theoretical bases of the rationales, and also the empirical failures of the preclinical phases that preceded the study.

Considering all these difficulties, various international actions have invested heavily in the qualification processes of scientific literature in health and related Information such the *Minimum* about Microarray areas. as a and the Minimum Information for Biological and *Experiment* (MIAME) *Biomedical Investigation*, (MIBBI)<sup>7.8</sup>. In research in the health area, uniformity requirements for manuscripts sent to biomedical journals can be considered one of the main actions in the search for the qualification of publications for authors, editors, analysts and scientific publishers<sup>9</sup>, as well as such as Enhancing the *Quality and Transparency of Health Research* (EQUATOR)<sup>7</sup>, which brings together researchers, editors, research methodology specialists and others interested in improving the quality and transparency of publications through guidelines that help improve the experimental aspects and results.

Considering that the evaluation of the methodological and information quality allows the analysis of the execution and application of a research <sup>3,6</sup> that can qualify scientific production, this article sought to identify the evaluation scales of the methodological quality of the scientific articles and checklists of the quality of information in research in the health area, based on their basic characteristics regarding the number of *items*, psychometric properties of validity and reliability, applications and limitations. **METHODS** 



This systematic review was conducted in accordance with the recommendations of the Cochrane Collaboration. <sup>10,11</sup> The studies were searched in the databases *Web of Science , Journals@ovid, Science Direct, Scopus, SportDiscus, Mary Ann Liebert* and *Oxford Journals Online* according to the descriptors of the *Medical Subject Headings* (MeSH) and Health Descriptors. (DECs): a) lists (*checklist*), b) scale (*scales*); c) critical appraisal of methodology, d) quality assessment, available in the keywords of the articles. The identification, manipulation and control of **REFERENCES** and files were carried out with *EndNote* (version 3.5).

#### **IDENTIFICATION OF STUDIES**

Two procedures were carried out to identify the studies: a) systematic review; b) bibliometric review. In the systematic review, studies were identified based on inclusion criteria that aimed to obtain complete articles, with indexed titles in the English language, produced between 2007 and 2012, coming from the health area. The exclusion criteria were used to discard articles that did not present sufficient information, mainly regarding their psychometric properties. The search was standardized and carried out by two independent reviewers (FCS and RS) who initially proceeded to read the titles, then the abstracts, and finally the full article. The articles recognized after comprehensive reading were recovered and in case of divergences in obtaining the studies, the procedures were inversely repeated by both reviewers until the discrepancies were corrected.



#### SELECTION OF STUDIES

After the identification and primary selection of eight articles, a bibliometric review was carried out, seeking to identify authors/reference works on the topics in question, and who did not have the time or language of publication established. With the manual search of the works available in the references of the selected articles, the existence of 51 studies on the use, development and psychometric properties of scales and lists for evaluating the methodological and information quality were identified. Of these 59 studies, 25 articles were selected because they dealt with the development of scales and lists for evaluating methodological and information quality (Fig. 1). The main reasons for exclusion in the systematic review and bibliometric review were: a) very specific studies that could not be applied in health areas; b) quality assurance studies; c) duplicate studies and d) studies of application, but not development, of scales and lists for evaluating methodological and information quality.

#### DATA EXTRACTION AND ANALYSIS

The extraction of information from the selected articles was aimed at obtaining the identification of the scales, the authors, the year of publication of the article and consequently the scale, the number of *items* in the scale, the indices and the implications of the validity of the scale, and the indices and implications of the reliability of the scales, as well as their applications and limitations.



The analysis of the information was carried out in a qualitative descriptive manner in relation to its implications and limits, and quantitatively in relation to the investigated indices. The article selection process is shown in Figure 1. The results on the identification of the scales and lists, authorship, year of publication, number of *items*, validity and reliability of the scales are presented synoptically through tables <u>1</u> and <u>2</u>.

### RESULTS

A total of 14 scales ( $\underline{table 1}$ ) and 11 lists ( $\underline{table 2}$ ) were identified, which offer different evaluation focuses and which demonstrate the advances in the use of tools for evaluating the methodological and information quality, as well as can offer benefits for current science, especially for the health area.

### SCALES

- *Jadad Scale* : <sup>12</sup> was originally developed and validated to independently assess the quality of RCTs on pain, but has been used for other purposes, including as a "gold standard." It presents a five-point quality score, with two additional points for appropriate randomization methods and secrecy of placement, ranging from 0 (weak) to 5 (good) <sup>.12</sup> The first *item* deals with the way patients are randomized; the second, from the use of the double-blind; and the third from the loss of individuals. <sup>12</sup> This scale presented evidence of concurrent validity and demonstrated strong correlation with various scales. The interclass correlation



coefficient (ICC) ranged from satisfactory to excellent and Kappa revealed weak to excellent reliability.

- *Maastricht Scale* : <sup>13</sup> methodologically evaluates the quality of a clinical trial and has an educational function regarding its conception and publication. It consists of 15 *items* based on methodological quality evaluation criteria, which are divided into 47 sub-*items* and total 100 points in three dimensions of the quality of a clinical trial: internal validity, external validity and statistical method. The scale includes four response options, and values are attributed to *items* that reflect relative importance. <sup>14</sup> The concurrent validity with the Delphi List and the Jadad scale presented a strong correlation and the ICC was excellent.

- *Single-Case Experimental Design Scale* (SCED): seeks to evaluate the methodological quality of case studies. <sup>15</sup> Yates seeks to measure the quality of RCTs of psychological interventions in the treatment of chronic pain and its use has been very limited. <sup>16,17</sup> The *SCED* was constructed including 11 *items*, of which 10 are used to evaluate the methodological quality and the use of statistical analysis. <sup>18</sup> The *SCED* shows an excellent level of interrater reliability when the total score is used. For all *items*, reliability ranged from satisfactory to excellent by two raters.

- *Van Tulder Scale* : <sup>19.20</sup> analyzes the threats to the validity of RCTs based on the elements of adequacy of the random method, concealment of treatment placement, poor vision and analysis by treatment intention. It is an 11-point scale that analyzes threats to validity and includes elements of methodological adequacy of the research. <sup>21,22</sup> This list barely includes the internal validity criteria. <sup>23</sup> Face validity and content validity were analyzed. This scale presented a strong correlation in



concurrent validity with other scales and in interobserver reliability, and showed reliability that ranged from weak to moderate.

- *PEDro Scale* : <sup>8.24</sup> was developed to be used in experimental studies. It offers an important source of information to support clinical evidence-based practice. This scale evaluates the internal validity and presentation of the statistical analysis of the studies. It presents 10 *items* on internal validity and presentation of the statistical analysis. The presence of indicators of the quality of the evidence presented is assigned 1 point and not 0 points. <sup>25</sup> This scale was validated, although the type has not yet been identified. Reliability varied from good to excellent in Kappa and from bad to excellent in ICC.

- *Bizzini Scale* : <sup>26</sup> seeks to evaluate the quality of RCTs on patellofemoral pain syndrome based on four main criteria (population, interventions, effect size, and data presentation and analysis) and 14 specific criteria. 25 points were attributed to each of the four main criteria for a total of 100 points, and a maximum of 5 to 10 points for the specific criteria. All criteria range from 0 to 5 or 0 to 10 points, with 0 for an inadequate description and the maximum number of points for a detailed and appropriate description. <sup>26</sup> Face validity and content validity were analyzed and the ICC ranged from satisfactory to excellent.

- *Chalmers Scale*, <sup>27</sup> evaluates quality through 32 *items*. The score evaluates two dimensions of quality (internal generalization and external validity) with maximum scores of 88 points. <sup>27-29</sup> Face and content validity were analyzed, and therefore this scale requires additional validation. In relation to the ICC, there was a satisfactory to excellent variation, and in the *test - retest* it was excellent.



- *Reisch Scale* : <sup>27.30</sup> evaluates the quality of RCTs on the use of aspirin in coronary heart disease. It was recently adapted and presented reliability for studies on pharmacological treatments for osteoporosis. <sup>29</sup> It is made up of 34 *items* divided into 13 domains (objectives, experimental project, sample determination, randomization and stratification, description and aptitude of the participants, use of comparison with the control group, procedures for treatment or management, concealment , reduction of participants, analysis and evaluation of participants in treatment, presentation and analysis of data, and recommendations and conclusions). Face and concurrent validity presented a strong correlation with other scales. The CCI was considered satisfactory.

- *Yates Scale* : <sup>16.17</sup> seeks to measure the quality of RCTs of psychological interventions in the treatment of chronic pain and its use has been very limited. It consists of 8 *items* and 26 sub *-items* and its use has been limited, since it was cited only once by the same group of authors. <sup>16</sup> To demonstrate the validity of the scale, the face, content and discrimination validity were analyzed. In this scale, the Kappa coefficient of each of the *items* analyzed varied from weak to satisfactory and the ICC was considered excellent.

- *Detsky Scale* : evaluates the quality of clinical trials of parenteral nutritional support for patients undergoing major surgery. <sup>23,31</sup> It consists of 13 variables with five main *items*, which reaches a maximum score of 14 points. In the concurrent validity carried out with the Reisch, Jadad and Van Tulder scales, a strong correlation was presented. Spearman correlation ranged from strong to excellent.

- *Sindhu Scale* : <sup>27</sup> is a tool for evaluating the methodological quality of primary RCTs to be included in a meta-analysis. It consists of 53 *items* subdivided into 15



dimensions on the methodological quality of RCTs used in meta-analysis.<sup>32</sup> In the reliability *test*, *the ICC demonstrated a strong correlation;* That is why it was evaluated with only two evaluators. The Sindhu presented high correlation coefficients for face, content, and criterion validity compared to the Chalmers scale, but still needs further testing.

- *Oxford Pain Validity Scale (OPV)*: it was established with the purpose of measuring the validity of the results of RCTs to allow the classification of trial results according to the validity of the evaluations. <sup>13</sup> The OPV is composed of five main *items*, and the last *item* is divided into four qualitative sub-*items*. <sup>33</sup> Only face validity was analyzed and reliability was not reported.

- *Arrivé Scale* : it was built with the objective of evaluating the methodological quality of clinical investigations that use radiological examinations. <sup>34</sup> It evaluates the methodological quality through 15 *items* related to the study design and the characteristics of the population, beyond a description of the image analysis. <sup>34</sup> Reliability was measured between two observers, and presented good to high agreement. The CCI showed high agreement. Only face validity was analyzed.

- *Newcastle-Otawa Scale (NOS)* : it was developed to evaluate the quality of nonrandomized studies, seeking to incorporate quality evaluations in the interpretation of meta-analysis of the results obtained. <sup>35,36</sup> The NOS evaluates quality based on content, design, and ease of use in the interpretation of the meta-analysis. It is composed of eight *items*, divided into three dimensions (comparison, selection, type of study) from cross-sectional, cross-sectional or case-control research. <sup>35,36</sup> Face and content validity were established based on a critical review



of the *items* by specialists in the area. <sup>36</sup> Reliability was not reported, but it was recently used in the meta-analysis that observed that the scale proves to be reliable and valid.

#### LISTS

- *Delphi List* : it represented the first step towards a registry for evaluating the quality of RCTs, <sup>18</sup> but it has not been used correctly, since it must be used together with other instruments for evaluating methodological quality. <sup>18</sup> This list evaluates RCTs through eight questions about the randomization method used, concealment of placement, blinding of the evaluator, therapist and patient, and statistical analysis. <sup>18,37</sup> This list presented greater validity compared to other scales, but internal consistency and validity still need to be established. Interobserver reliability ranged from good to excellent and presented strong ICC.

- *Maastricht Amsterdam List (MAL)* : it is recommended for systematic reviews that investigate spine problems, specifically low back pain. <sup>14</sup> The MAL is composed of 19 *items* that consist of the evaluation of internal validity, the descriptive criteria and the statistical aspects adopted. The general numerical score adopted is from 0 to 19 points. <sup>14,33,38</sup> Face and content validity were analyzed and inter-observer reliability reported ranged from weak to good.

- *MOOSE Registry*: is a detailed checklist to report the meta-analysis of observational studies in epidemiology. <sup>39</sup> The MOOSE list is organized into 35 dichotomous *items* and divided into six sections that address information about the



Abstract, research strategy, methods, results, discussion and conclusions. <sup>39</sup> The validity and reliability of this list were not identified in the studies analyzed.

- *Andrew List* : evaluates the quality of clinical trial-type studies that use X-ray contrast media. <sup>40</sup> The Andrew <sup>40</sup> was created in the 1980s and modified in the 1990s by the same authors and is composed of 11 *items* that seek to evaluate the quality of clinical trial-type studies. The modified List presented weak inter-observer reliability.

- *Downs and Black List* : it is used for randomized and non-randomized studies, <sup>41</sup> and was recently reviewed for use in evaluating the quality of epidemiological studies, <sup>42</sup> but it is not yet applicable for prevalence studies. <sup>43</sup> It consists of 27 *items* with five subscales (registration, external validity, errors, confusion, power) and can be used in various types of studies. <sup>41</sup> The quality index was highly correlated with the Trial Group Registry Register scores. The internal consistency was considered adequate, as was the ICC. This list was recently expanded by two new criteria and validated for use in epidemiological studies. <sup>43</sup>

- *Nguyen List* : was developed to evaluate the methodological quality of the studies. <sup>44,45</sup> The Nguyen <sup>45</sup> was developed with attribution for each of the 18 items of a numerical score in accordance with the accompanying guidelines, but its use is currently not recommended. Face validity was analyzed, and reliability was not reported. With all this, some of the *items* on the list must be evaluated subjectively, since they are not directly related to the quality of the study. <sup>44</sup>

- *CONSORT Statement* : it was developed by a group of scientists and editors for critical evaluation and interpretation of RCTs, <sup>26</sup> in search of more precise and reproducible evaluations. <sup>21,46-48</sup> It was first published in 1996 and updated in 2001.



In January 2007, the CONSORT Statement was subsequently modified for a journal and was published as the CONSORT Statement 2010. <sup>47</sup> The CONSORT presents 25 *items*. <sup>49</sup> The questions are answered with yes (1 point) or no (0 points), for a maximum of 25 points. The CONSORT presented 94% agreement between reviewers. <sup>twenty-one</sup>

- *COCHRANE Collaboration Criteria* : <sup>11</sup> serve to evaluate the patient's allocation status and classify studies as adequate, doubtful, inadequate and not performed. For the COCHRANE criteria, adequate randomization is necessary to generate an error-free comparison between groups. <sup>50</sup> Validity and reliability were not reported.

- *STROBE Statement:* seeks the quality of information from observational studies with a focus on prevalence (sectional, case-control, cross-sectional), <sup>51</sup> and serves as support for editors and reviewers. <sup>52</sup> It consists of 22 *items* on the title of articles, Abstract, introduction, methods, results, discussion sections and other information. A total of 18 *items* are common to the three designs; As for the other *items*, they are specific to the design. For some *items*, information must be given separately for cases and controls in case-control studies, or exposed and non-exposed groups in the cross-sectional study and cross-sectional studies. <sup>52</sup>

- AMSTAR: is an instrument that seeks to evaluate the methodological quality of systematic reviews of RCTs. <sup>16,32</sup> It is composed of 11 *items* of reliable and valid measures for evaluating the methodological quality of systematic reviews of RCTs. <sup>53,54</sup> Face, content and construct validity were tested. The CCI was tested in relation to other scales. For reliability, the ICC for the total score was



excellent. However, additional studies focusing on reproducibility and construct validity are still necessary. <sup>54</sup>

- *PRISMA* : is a checklist that aims to seek the transparency of information from systematic reviews important to the scientific qualification process of these studies. It has 27 *items* and a four-phase flow chart that includes *items* considered essential for the transparent communication of a systematic review. <sup>55</sup> Validity and reliability were not reported.

Synoptically, it can be stated that scales and checklists currently have a wide range of applications, but also many limitations, since they can be used in the evaluation of studies of different types, for different populations and on different health approaches, such as Jadad, Maastricht, Van Tulder, Bizzini, Chalmers, Yates, Detsky, Sindhu, OPV, Delphi, Andrew Modified, Downs and Black, CONSORT, Cochrane Criteria and AMSTAR in RCTs; Reisch, Detsky, NOS, Andrew Modified and Downs and Black in non-randomized controlled trials; MAL, Detsky, Sindhu, PRISMA, Critérios Cochrane and AMSTAR in systematic reviews; the Sindhu, NOS, MOOSE and Cochrane Criteria in the metaanalyses; the PEDro and the Cochrane Criteria in experimental studies; MOOSE, Downs and Black and STROBE in epidemiological studies; the SCED and PRISMA in clinical research; the Arrivé and the Andrew Modificado in specific clinical investigations. It can also be verified that some scales and lists are no longer recommended, such as the Nguyen, and must be used in conjunction with other scales and lists, such as the Delphi.



It can also be said that synoptically the scales and checklists identified present a wide variety of validity and reliability indices. The scales and lists that present adequate validity are: Jadad, Delphi, Maastricht, MAL, Van Tulder, Bizzini, Chalmers, Reisch, SCED, Andrew Modified, Yates, Detsky, CONSORT, Sindhu, Downs and Black, Nguyen, OPV, Arrivé, Cochrane, AMSTAR and NOS criteria. The scales and lists that presented adequate reliability are: Jadad, Delphi, Maastricht, MAL, Van Tulder, PEDro, Bizzini, Chalmers, Reisch, SCED, Andrew Modified, Yates, Detsky, CONSORT, Sindhu, Downs and Black, Arrivé, Cochrane Criteria , AMSTAR, PRISMA and NOS. Some scales and lists were either not developed or their validity (PEDro, STROBE, PRISMA) and reliability (MOOSE, Nguyen, OPV, STROBE) of the processes were not reported, or they also did not present adequate indices.

### DISCUSSION

Considering the results presented, the appearance of 14 scales (including their modifications) and 11 lists was confirmed. With all this, it can be verified that only 14 stops (Jadad, Maastricht, PEDro, Van Tulder, Bizzini, Chalmers, Reisch, SCED, Andrew Modifica, Yates, Detsky, Sindhu, Downs and Black and Arrivé) and 4 lists (Delphi, MAL, CONSORT and AMSTAR) presented duly investigated psychometric properties (validity and reliability). It can be identified that the Jadad scale presented the best evidence of validity and reliability, because it was tested in different contexts, and—in conjunction with the Delphi List—there is evidence of greater validity compared to the other scales and lists (MAL, Van Tulder, PEDro and Bizzini). Meanwhile, the Delphi List lacks internal consistency and construct



validity. These psychometric properties are relevant because they indicate that the construct, in this case the methodological quality, is fully represented by the *items* of the scale (internal consistency), and that the scores of a scale must be adequate based on pre-defined hypotheses. <sup>5,56</sup> The MOOSE, Cochrane Criteria, STROBE and PRISMA lists did not present psychometric properties, and the Nguyen, Oxford and NOS scales only reported validity.

Specifically in the case of the Jadad scale, it can be found that although it has been developed and validated to evaluate the quality of studies carried out on pain, it has also been used extensively in other clinical areas.<sup>22</sup> Currently, countless clinical trials include the *items* of the Jadad scale in their methodology in order to carry out study with good methodological quality. In this sense, Herbison and a others <sup>57</sup> concluded that the Jadad scale may not be sensitive or sufficient to distinguish between different levels of quality. Therefore, the use of the Jadad scale and its validity must be reevaluated for different areas of research. In the case of the Delphi list, it can be verified that despite having been built specifically for the evaluation of the quality of RCTs, <sup>18</sup> it has been used in various other areas. One of the factors that contribute to this expansion of its use is related to the fact that the psychometric properties of the Delphi list indicate that the methodological quality is fully represented by the *items* of the scale (internal consistency), and that the scores of a scale are based on predefined hypotheses, <sup>16</sup> which strongly led to the replacement of the Maastricht scale that was widely used (MAL, Van Tulder, PEDro). <sup>40</sup> The modified Maastricht scale, developed from a valid and reliable scale, cannot be considered valid and reliable until it is tested. According to Streiner and Norman, <sup>56</sup> modifications to existing



scales often require new validity studies. This means that the psychometric properties of the modified scale have to be evaluated to ensure that the new scale can truly identify good or poor methodological quality.

Generally, it can be verified that the existing scales and lists provide methodological resources for readers, authors, reviewers and scientific editors based on different proposals and objectives, and are themselves an element of constant review, such as the PEDro scale, a modification of the Delphi list, which offers a more comprehensive measure of the methodological quality of the poststroke rehabilitation literature the Jadad compared to scale, <sup>58</sup> the *Downs* and *Black* list, which was reviewed for the evaluation of quality of the population based on epidemiological studies, 42 but it is not yet applicable for prevalence studies, <sup>43</sup> and even the NOS, which despite being in the process of evaluating the validity of the scale, but already with indications of being reliable and valid, <sup>25</sup> have problematic *items*, with lack of adequacy of the analysis, lack of information related to reliability and validity.<sup>29</sup>

The scientific production of various countries has grown considerably in recent decades. <sup>59,60</sup> In this sense, the evaluation of the quality of the studies becomes essential due to their transparency, visibility, rigor and impact of scientific production and publication. Editors, reviewers and researchers must increasingly have knowledge of the tools for evaluating methodological quality and information, because the standards for the publication of articles are becoming increasingly demanding and strict. With all this, the necessary criticism of these processes of qualification of scientific production must be highlighted, because it must be considered that the majority of the revised scales and lists do not meet the



methodological the standards, and requires inclusion of appropriate important *items* that must be evaluated in terms of the type of study, application, psychometric capacity, mainly validity and reliability, among others. In this sense, researchers, editors and analysts must take special care and attention when using certain scales to evaluate the quality of scientific studies, because the limitations information presented must be interpreted with caution. According and to Serra,<sup>7</sup> the need to publish results in a clear and transparent manner positively influences the formation of new knowledge, and consequently increases confidence in the conclusions when the study is carried out with adequate methodological rigor. Therefore, despite the aforementioned care, the spread of the use of methodological and information evaluation tools increasingly qualifies and legitimizes scientific production, mainly in the area of health.

It is concluded that, considering the increase in research linked to health sciences and related areas, it becomes essential to establish an evaluation of the methodological quality of the studies carried out since this procedure can contribute to avoiding inconsistent publications, and to improve the literature selection process in terms of its validity, relevance and clinical applicability. In this way, with the intention of ensuring the scientific rigor of the research based on the results of this systematic review, it can be assured that many scales are being produced to evaluate the methodological quality of scientific articles and information quality checklists. in research in the health area. It can be highlighted that the scales and lists differ from each other in relation to the number of *items*, validity, reliability and scoring parameters, and that several of these present valid and reliable psychometric properties, beyond the fact that these scales and lists are



applicable. to empirical studies, mainly randomized clinical trials, and to reviewal studies, mainly meta-analytical ones, both in the search for methodological quality and the quality of the information. Therefore, it is worth noting that the use of methodological and information evaluations is greater, positively influencing the increase in scientific production, especially in the area of health and related areas.

#### REFERENCES

1. Verhagen AP, de Vet HCW, de Bie RA, Boers M, van den Brandt PA. The art of quality assessment of RCTs included in systematic reviews. J Clin Epidemiol. 2001;54(7). p. 651-4.

2. Emerson JD, Burdick E, Mosteller F, Chalmers TC. An empirical study of the possible relation of treayment differences to quality scores controlled randomised clinical trials. Contr Clin Trials. 1990;11:339-52.

3. Juni P, Altman DG, Egger M. Systematic reviews in health care - Assessing the quality of controlled clinical trials. Brit Med J. 2001;323(7303). pp. 42-6.

4. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials - an annotated-bibliography of scales and checklists. Contr Clin Trials. 1995;16(1):62-73.

5. Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. Oxford: 2003.

6. Greenfield M, Rosenberg AL, O'Reilly M, Shanks AM, Sliwinski MJ, Nauss MD. The quality of randomized controlled trials in major anesthesiology journals. Anesth Analg. 2005;100(6):1759-64.



7. Serra LC. Good publication practices, their evolution and the expected impact on public health. Rev Cubana Sal Públ. 2012;38:725-33.

8. Taylor CF, Dawn F, Sansone SA, Aerts J, Apweiler R, Ashburner M, et al. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nat Biotech. 2008;26:889-96.

9. Munhoz Junior E. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publications. Epidemiol Serv Saúde. 2006;15:7-34.

10. Higgins JPT, Green S. Cochrane Handbook of Systematic Reviews of Interventions. The Cochrane Collaboration; 2011.

11. Higgins JPT. Assessing risk of bias in included studies. The Cochrane Collaboration; 2005.

12. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Contr ClinTrials. 1996;17(1):1-12.

13. de Vet HC, de Bie RA, van der Heijden GJ, Verhagen AP, Sijpkes P, Knipschild PG. Systematic reviews on the basis of methodological criteria. Physioth. 1997;83(6):284-89.

14. Van Tulder MW, Assendelft WJJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration back review group for spinal disorders. Spine. 1997;22(20). p. 2323-30.

15. Tate AR, Gennings C, Hoffman RA, Strittmatter AP, Retchin SM. Effects of bone-conducted music on swimming performance. J Strength Cond Res. 2012;26(4):982-8.



16. Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess the quality of randomized controlled trials: A systematic review. Physioth. 2008;88(2):156-75.

17. Yates SL, Morley S, Eccleston C, Williams ACD. A scale for rating the quality of psychological trials for pain. Pain. 2005;117(3):314-25.

18. Verhagen AP, de Vet HCW, de Bie RA, Kessels AGH, Boers M, Bouter LM, et al. The delphi list: A criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. J Clin Epidemiol. 1998;51(12). p. 1235-41.

19. Sung L, Nathan PC, Lange B, Beyene J, Buchanan GR. Prophylactic granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor decrease febrile neutropenia after chemotherapy in children with cancer: A metaanalysis of randomized controlled trials. J Clin Oncol. 2004;22(16). p. 3350-6.

20. Dekker A, Bulley S, Beyene J, Dupuis LL, Doyle JJ, Sung L. Meta-analysis of randomized controlled trials of prophylactic granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor after autologous and allogeneic stem cell transplantation. J Clin Oncol. 2006;24(33). p. 5207-15.

21. Warschkow R, Tarantino I, Jensen K, Beutner U, Clerici T, Schmied BM, et al. Bilateral Superficial Cervical Plexus Block in Combination with General Anesthesia Has a Low Efficacy in Thyroid Surgery: A Metaanalysis of Randomized Controlled Trials. Thyroid. 2012;22(1):44-52.



22. Clark HD, Wells GA, Huet C, McAlister FA, Salmi LR, Fergusson D, et al. Assessing the quality of randomized trials: Reliability of the Jadad scale. Contr Clin Trials. 1999;20(5):448-52.

23. van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board Cochrane C. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. Spine. 2003;28(12):1290-9.

24. Sherrington C, Herbert RD, Maher CG, Moseley AM. PEDro. A database of randomized trials and systematic reviews in physiotherapy. ManTher. 2000;5(4):223-6.

25. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther. 2003;83(8):713-21.

26. Bizzini M, Childs JD, Piva SR. Systematic review of the quality of randomized controlled trials for patellofemoral pain syndrome. Journal of Orthopaedic & Sports Phys Ther. 2003;33(1):4-20.

27. Chalmers TC, Jr Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, et al. A method for assessing the quality of a randomized control trial. In: Smith HJR BB, editor. Contr Clin Trials.1981;2(1):31-49.

28. Sculier JP, Berghmans T, Castaigne C, Luce S, Sotiriou C, Vermylen P, et al. Maintenance chemotherapy for small cell lung cancer: a critical review of the literature. Lung Canc. 1998;19(2):141-51.

29. Berard A, Andreu N, Tetrault JP, Niyonsenga T, Myhal D. Reliability of Chalmers scale to assess quality in meta-analyses on pharmacological treatments for osteoporosis. An Epidemiol. 2000;10(8):498-503.



30. Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. Pediatr. 1989;84(5):815-27.

31. Detsky AS, Naylor CD, Orourke K, McGeer AJ, Labbe KA. Incorporating variations in the quality of individual randomized trials into metaanalysis. J Clin Epidemiol. 1992;45(3):255-65.

32. Sindhu F, Carpenter L, Seers K. Development of a tool to rate the quality assessment of randomized controlled trials using a Delphi technique. J Adv Nurs. 1997;25(6):1262-8.

33. Smith LA, Oldman AD, McQuay HJ, Moore RA. Teasing apart quality and validity in systematic reviews: an example from acupuncture trials in chronic neck and back pain. Pain. 2000;86(1-2):119-32.

34. Arrive L, Renard R, Carrat F, Belkacem A, Dahan H, Le Hir P, et al. A scale of methodological quality for clinical studies of radiologic examinations. Radiology. 2000;217(1):69-74.

35. Aarts JWM, van den Haak P, Nelen WLDM, Tuil WS, Faber MJ, Kremer JAM. Patient-focused Internet interventions in reproductive medicine: a scoping review. Hum Reprod Upd. 2012;18(2):211-27.

36. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in metaanalyses. 2000.

37. Verhagen AP, de Vet HCW, de Bie RA, Kessels AGH, Boers M, Knipschild PG. Balneotherapy and quality assessment: Interobserver reliability of the Maastricht criteria list and the need for blinded quality assessment. J Clin Epidemiol. 1998;51(4). p. 335-41.



38. Seferiadis A, Rosenfeld M, Gunnarsson R. A review of treatment interventions in whiplash-associated disorders. Europ Sp J. 2004;13(5):387-97.

39. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Metaanalysis of observational studies in epidemiology: a proposal for reporting. J Am Med Assoc. 2000;283(15). p. 2008-12.

40. Andrew E, Eide H, Fuglerud P, Hagen EK, Kristoffersen DT, Lambrechts M, et al. Publications on clinical trials with X-ray contrast media: differences in quality between journals and decades. Europ J Radiol. 1990;10(2):92-7.

41. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Comm Heal. 1998;52(6). p. 377-84.

42. Macfarlane TV, Glenny AM, Worthington HV. Systematic review of population-based epidemiological studies of oro-facial pain. J Dent. 2001;29(7):451-67.

43. Shamliyan T, Kane RL, Jansen S. Quality of Systematic Reviews of Observational Nontherapeutic Studies. Prev Chron Dis. 2010;7(6). p. 9-195.

44. Giannakopoulos NN, Rammelsberg P, Eberhard L, Schmitter M. A new instrument for assessing the quality of studies on prevalence. Clin Oral Invest. 2012;16(3). p. 781-8.

45. Nguyen QV, Bezemer PD, Habets L, Prahl-Andersen B. A systematic review of the relationship between overjet size and traumatic dental injuries. Europ J Orthod. 1999;21(5):503-15.



46. Huwiler-Muntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. J Am Med Assoc. 2002;287(21):2801-4.

47. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. 2012;65(3). p. 239-46.

48. Moher D, Jones A, Lepage L, Grp C. Use of the CONSORT statement and quality of reports of randomized trials. A comparative before and after evaluation. J Am Med Assoc. 2001;285(15):1992-5.

49. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Springer Verlag. 2003;7(1). pp. 2-7.

50. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical-evidence of bias - dimensions of methodological quality associated with estimates of treatment effects in controlled trials. J Am Med Assoc. 1995;273(5):408-12.

51. von Elm E, Altman DG, Pocock SJ, Gotzsche PC, Vandenbroucke JP, Initiative S. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Brit Med J. 2007;335(7624):806-8.

52. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Explanation and Elaboration. Epidemiol. 2007;18(6):805-35.



53. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013-20.

54. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. Bmc Med Res Methodol. 2007;7(1):10.

55. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and metaanalyses of studies that evaluate healthcare interventions: explanation and elaboration. Brit Med J. 2009;(6):1-27.

56. Streiner DL, Norman GR. Validity. Health measurement scales: A practical guide to their development and use: Oxford; 2004. p. 172-93.

57. Herbison P, Hay-Smith J, Gillespie WJ. Adjustment of meta-analyses on the basis of quality scores should be abandoned. J Clin Epidemiol. 2006;59(12):1249-56.

58. Bhogal SK, Teasell RW, Foley NC, Speechley MR. The PEDro scale provides a more comprehensive measure of methodological quality than the Jadad Scale in stroke rehabilitation literature. J Clin Epidemiol. 2005;58(7):668-73.

59. Nunes ED. A review of research studies conducted on scientific production in collective health in Brazil. Scientometrics. 1999;44(2):157-67.



# Editorial Peer Review: Rejection Of The Manuscript, Deficiencies In The Review Process, Systems For Its Management And Use As A Scientific Indicator

Victor E. Nzube, Faculty of Agriculture, University of Nigeria, Nigeria

### ABSTRACT

Digital platforms and the dynamics and evolution of scientific journals have allowed the development of various models of the editorial peer review process for the evaluation of scientific manuscripts prior to their publication. This article continues the analysis of peer review, with emphasis on the management of rejected articles, the appointment of reviewers, the main deficiencies of peer review (based on the assignment of reviewers, the performance of author roles, reviewer and editor, and attempts to alleviate deficiencies in the process), systems for online management and the use of peer review as an indicator of research performance. All these topics are analyzed in the context of science systems and communities, their impact on citation, and to facilitate their possible integration for requirements according the of practical purposes to each journal. **Keywords:** peer review for publication, peer review of research, academic review, journal article, electronic publication, scientific and technical publications.



### INTRODUCTION

The editorial peer review (PR) process is vital to ensure and control the quality of articles prior to publication, <sup>1</sup> and like any human work, it is subject to distortions and improvement. Its satisfactory result, the acceptance of the manuscript, does not generate as many questions as its rejection, the latter influenced by several elements, among them those that concern the quality of the PR. This article analyzes the rejection of manuscripts and the PR models used to manage articles rejected due to PR deficiencies, as well as their main limitations related to speed, selection of each role, considerations about the stages of the process and their control of quality. Online management systems and the implications of publishing reviews are also discussed.

#### **REJECTION OF THE MANUSCRIPT AND ITS INTERPRETATION**

The editorial rejection of manuscripts after PR has historically been considered a quality criterion for scientific journals, since the PR process is considered efficient when it detects and discards those manuscripts sent to journals that do not meet their quality requirements and the good research practices. <sup>2</sup> It is considered to a certain extent predictive of high rates of methodological quality of the articles and is linked to the most demanding criteria of the thematic leaders who make up the editorial board of the journals. <sup>3</sup> It has always been assumed to invalidate subsequent resubmissions when it is due to deficiencies in the scientific procedure or due to non-compliance with the general parameters of scientific communication, even at the level of publication societies such as the case of the journals of the American Association for Microbiology (ASM). <sup>4</sup> This does not seem to be the



trend among articles rejected in first-tier journals, which are explicitly considered possible for reconsideration by second-tier journals.<sup>5</sup>

The rejection rate of a journal is commonly expressed as the rejection rate (in percentage) or its counterpart as the acceptance rate. A very high rejection rate has been considered counterproductive to the rapid dissemination of science when it is due to factors unrelated to the quality of the research evaluated. However, this index is among the criteria that guarantee the prestige of a publication. <sup>6</sup> In toplevel journals, it denotes the selection through PR of results with the greatest potential impact, given the avalanche of manuscripts to be published<sup>7</sup> and indicates which articles may be relevant to readers.<sup>6</sup> The appearance of the open process PR model made it possible to favorably modify the nature of this indicator, since in this model the authors generally send articles with sufficient quality so that the probability of acceptance is high after the PR, <sup>6</sup> given the nature review public. The manuscript can be rejected for multiple factors, <sup>8</sup> that involve all the roles of the process (author-editor-reviewer) and even the scientific community from which they come. During submission, information is requested from the authors about any previous submission and review of the manuscript in another journal. However, not all authors refer to it or pay attention to the elements indicated in them, so that its subsequent publication does not imply the deficiencies initially detected or duplicate the effort.<sup>9</sup> Although it is true that authors do not communicate previous submissions mainly due to the possible negative perception about the quality of the manuscript based on a previous rejection, it is already allowed to transfer these reviews between journals, <sup>10</sup> mainly when resubmissions



are arranged between the editorial boards. of the journal that rejects and the one that will reevaluate the manuscript.

In a very recent study by *Calcagno* and others, it was shown that, in fact, articles that suffer at least one rejection prior to publication tend to be cited more than those accepted on a first attempt.<sup>11</sup> These authors used a new approach to the dynamics of the sending and forwarding flows of scientific articles, by studying surveys of authors of published articles. The results should be considered with caution since the sample included 25% of articles rejected in their first submission, so it is necessary to wait for such a strategy and its interpretation to be adequately validated. However, it is curious that if the manuscript was ever rejected during its publication process, resubmission would increase the number of reviewers, so a greater number of reviewers would better evaluate the possible interest for the community of readers (interested public). For this reason, resubmission works in a similar way to post-PR and adds value to reporting the previous rejection whenever it is due to scientific controversy or space limitations to publish in the journal. Thus, rejection would no longer be considered a limitation, unless it is related to deficiencies in the content and investigative procedures. The publication of this data by the publishers and its reference by the authors during resubmission and subsequent publication would also be favored, perhaps under the first submission identifier and rejection type classifier to avoid incorrect assignment when the initial metadata of the article varies. All this forces us to reconsider the interpretations of rejection prior to effective publication, and opens a much-needed field of research to predict to what extent rejection by top-level journals could be a



predictive indicator of the impact of the manuscript, as well as its acceptance. and subsequent summons.

It should be noted that top-level publishing consortia have evolved to rescue within the same group of publications the effort invested in manuscripts rejected in their main journals due to space limitations or potential impact. For this they have enabled spaces to which these manuscripts are forwarded and published after a new PR. For example, the *Nature Publishing Group (NPG) consortium created the journal Scientific Reports* for this purpose, under an open access editorial management model.<sup>12</sup>

Variants of PR are mentioned below that allow us to editorially rescue those articles rejected due to distortion in the quality of the PR and not due to deficiencies in the quality or presentation of the research. These variants facilitate the forwarding of the manuscript between journals of the same thematic context, together with the information from the previous review; or the custody of the manuscript to avoid bias in cases where it addresses infrequent or very underrepresented topics in scientific communities.

#### MANAGEMENT MODELS FOR REJECTED ITEMS

There are two PR models that handle the resubmission of rejected articles more efficiently. They are *rebound*  $PR^{13}$  and *cascade* PR, also called *consortium* PR.<sup>14,15</sup> The first is in the same magazine, and the second between magazines linked thematically or structurally at the editorial level.



Rebound *PR*<sup>13</sup> preferentially addresses those works whose authors consider that the conventional, simply blind process was unfair or biased for any reason, unrelated to the scientific validity of the work. It was only put into practice by the journal *Antioxidants and Redox Signaling* in 2012, and allows authors of articles rejected due to intrinsic deficiencies of PR to suggest expert reviewers to re-evaluate the manuscript in the same journal, under the *open PR model per publication*. and the initial ruling can be reversed if the claim is admissible. This is a variant, as there may be others, of combining conventional and open models to this, such as the *British Journal of Sports Medicine*, which implemented a robust appeals system <sup>16</sup> in which editors involved in the first rejection are excluded from the appeals process.

*For its part, the cascade PR* model establishes the transfer of rejected articles between journals structured in editorial groups or consortia on the same topic. The manuscripts are transferred together with the reviews of the journal that rejected them. It is established by large publishing groups such as NPG <sup>17</sup>, the European Molecular Biology Organization (EMBO) <sup>14</sup> and the Neuroscience Peer Review Consortium (NPRC). <sup>15</sup> In the latter, the data referring to the initial submission that was rejected is hidden, since "the only information that the second journal will receive is the text of the reviews that were sent to the authors and, if the reviewers agree, the names of the reviewers (which are not communicated to the authors)". <sup>18</sup> On the other hand, this resubmission management model helps to increase the review standards of the second journal (which is generally of lower impact), 'contaminating' it with the review criteria of a higher impact one, while



managing the level of rejections generated by top-level journals. It is worth noting that *cascade PR* would be very useful if it were implemented at the national level by thematic field, especially in developing countries, or at the level of technology parks. At the same time, it is a variant that fixes to a certain extent the journal to which it is forwarded, as opposed to studying which one it was forwarded to spontaneously by the authors.

The development of a bibliometric indicator that adds value to the article/journal based on the journal to which it has been forwarded has recently been proposed. <sup>11</sup> In this sense, *cascade PR* could be the most likely niche to analyze the link between citation and resubmission, by setting the variable of hierarchical position between journals. In this way, each journal could be repositioned in the hierarchical order within the cascade, perhaps independently of the impact factor (which is potentially predictive) and based on the actual citation of the articles, which could progressively reconfigure its structure and dynamics.

It is worth highlighting that thematic redirection strategies are independent of rejection, and can be found in the instructions to the authors of journals, such as *Journal of Bacteriology*.<sup>4</sup> Even in journals structured *in cascading PR* it may involve forwarding to those that are both above and below the rejecting journal in position and relevance.

A strategy has also been implemented to avoid rejection of articles on controversial topics, in this case in the journal P *roceedings of the National Academy of Sciences USA*, <sup>19</sup> called Prearranged *editors*. In this, the author requests a member of that academy to follow up on the manuscript during the submission process, before



sending the manuscript, a designation that is carried out by the editorial office. This variant only applies if the manuscript to be considered "falls into a thematic area that is under-represented in the conclave, if the topic could be considered contrary to the prevailing view or if it is ahead of its time enough to not be given due attention." attention". <sup>19</sup> Once the consent and competence of said editor to assume the role of Concerted Editor is verified, his designation is communicated at the beginning of the PR.

### CONSIDERATIONS ABOUT THE PR PROCESS

#### **Manuscript Public**

The PR process is applied before the number of public interested in the manuscript is known (according to the conventional closed model, see below), an aspect that will only be known once it is published. Although the evaluation is more linked to the decision to publish, reviewers also assess the manuscript's potential to attract public attention. *In this sense, new platforms, in the form of discussion forums, allow articles to be enriched, or published as discussion notes (see, for example, the discussion forum related to Ploegh* 's article)<sup>20</sup> and hyperlinked as related information.

In certain analyzes it has been identified that PR presents difficulties in distinguishing research with the highest impact from average research, as well as bibliometric indicators based on citation analysis, <sup>21</sup> and that there is not necessarily a correlation between the quality of PR and the Subsequent citation



analysis to predict article impact. However, evidence has been found that PR, in fact, significantly selects articles that have high impact according to their number of citations. This was revealed when comparing the number of articles accepted in a high-impact journal with the number of articles rejected by the same journal and subsequently published in others in the same year.<sup>22</sup>

### Celerity

Speed in review is also vital, especially in those areas of intense experimental and high-impact competition. For example, in Chemistry, already in 1998 it was proposed that 80 days was the standard in the journal *Chemical Communications* and the intention was to reduce it. <sup>23</sup>

It has been proposed that in thematic areas of very active research it would be advisable to make the works available to the potentially interested public immediately after receiving them and that the editor of the journal be the one to rule, once the appropriate dissemination tools are available. <sup>24</sup>

Review time has been estimated to last approximately 5% of the entire editorial process. <sup>25</sup> However, *Loonen* and others found that the periods from submission to acceptance can range from 3 weeks to 15 months (also depending on the type of article and the topic in question). <sup>25</sup> The editorial policies of each editorial board and their actions to maintain process deadlines are decisive in shortening these periods.

DEFICIENCIES



It is recognized that there are failures during RP; for example, that it slows down the application processes for research projects and that it must be applied wisely. <sup>26</sup> Below are some considerations on the subject for a more exhaustive approach based on statistical studies. <sup>27.28</sup>

#### ASSIGNMENT OF REVIEWERS

The composition of the reviewers must be balanced. As *Schroter* et al. initially demonstrated, <sup>29</sup> reviewers suggested by authors tend to make more lenient evaluations than those chosen by the editor in *conventional PR*. Such an effect was also recently demonstrated for the *two-stage open RP* model , by studying 552 manuscripts processed in the journal *Atmospheric Chemistry and Physics*. <sup>30</sup> For this reason, it has been recommended to maintain a balance between the reviewers suggested by the authors and those suggested by the editors. Thematic affinity does not influence this, as it is the primary criterion for assigning reviewers both in *open PR* and co-reviewers in *conventional PR*.

A more recent study demonstrated that under real-world conditions in which reviewers are invited, PR tends to be systematically biased toward locating problems in experimental quality but without generating negative effects. <sup>31</sup> This tendency was explained as a reflection of cognitive conditioning as established by the logical reasoning processes for decision making (implicit *vs.* analytical), due to the priority given to experimental performance over other aspects when there are other more tasks. pressing and diverting attention at the time the review is assigned



(e.g., work overload). In this case and as explained by *Kadar*, <sup>31</sup> implicit mental processing depends, among other aspects, on pattern recognition, parallel reasoning, prior knowledge, requires little use of memory and generates results in a short time. In contrast, analytical processing is slow and sequential, requires considerable effort, depends on structured reasoning, and is limited by memory capacity. This could be the logical basis for simultaneous formal review and methodological evaluation of the manuscript, despite the fact that it was previously seen as a radical measure in the face of formal deficiencies. <sup>32</sup>

In Nursing, for example, it was identified during an editorial survey of authors and editors of three academic journals that authors inexperienced in the publication process (authors with 5 or fewer previous articles, about a third of the authors surveyed) were less satisfied with review by experienced authors. <sup>33</sup> In addition, there was special susceptibility to vague narrative revisions that did not provide clear guidelines to perfect the manuscript. This aspect is vital when the ruling is rejection.

Along these lines, an anonymous study of the perception of ethical problems during PR detected incompetent review as the most frequent problem of five main aspects (61.8% representativeness). <sup>34</sup> That the reviewers have due competence and are truly peers for the review act and during this is the net responsibility of the editorial board, when designating them in the *conventional PR* models , as well as the technical validity and relevance of the evaluations chosen in *open process PR* models . It is also pertinent to the authors when they propose reviewers. If the competency of reviewers is not adequately verified, PR suffers. It is decisive in the



quality of the PR that the editorial boards ensure compliance with the quality parameters of the review act.

It has been found that the training of the reviewers does not significantly influence the results of the PR.<sup>27</sup> The continuous inclusion of new reviewers does not weaken it, although there is insufficient empirical support for its use as a mechanism to guarantee the quality of biomedical research.

### ROLES OF EDITOR, REVIEWER AND AUTHOR

Peer review is learned through practice, and can be implemented as a learning method from the undergraduate stage, even with periods similar to those used in editorial practice. <sup>35</sup> To this end, initiatives have been implemented to train future researchers in peer review, for example, through undergraduate thesis evaluation protocols (BioTAP). <sup>36</sup> Students in the program increased their article writing and critical thinking skills, focused on identifying the audience need in each part of the text, and trained in the basic standards of scientific writing. This practice can be extended to publishers, with an overall profit in the process. <sup>37</sup> Emphasis has also been placed on training the authors on how to respond to the accusations, <sup>38</sup> in order to avoid rejections due to formal elements and deficiencies in communication.

During PR, editors cannot account for larger errors. Several studies have detected a lower tendency to detect fundamental errors by the reviewers who accepted the manuscripts. <sup>39,40</sup> Here the notion is reinforced that the editor must also be part of the act of evaluating the manuscript as another peer, to ensure that the PR has been



carried out with quality and to modulate communication. The editor could identify larger deficiencies and suggest to the referees that they address them directionally, once they have given their initial recommendation and if they have not taken them into account.

It has also been shown that the quality of reviews varies not only between reviewers when evaluating an article, but also changes over time for the same reviewer, and evaluation criteria tend to slowly deteriorate. <sup>41</sup> In a sample of 1,499 reviewers who worked for the *Annals of Emergency Medicine* for an average of 51.3 months over a 14-year period, *Callaham* and *McCulloch* identified that the quality of the first four review acts is predictive of the evolution of the reviewer, and that the inclusion of new reviewers is the group palliative to the deterioration of the predecessors. <sup>41</sup>

In the study by *Resnik* et al., <sup>34</sup> in addition to the incompetent review, the following aspects were identified in order of relevance within the deficiencies of PR: bias, request for inclusion of unnecessary references, comments from reviewers with personal attacks, and delay of publication to publish earlier articles by editors or reviewers on the same topic. This last aspect has been identified in certain highly competitive topics with research standards that evolve rapidly (for example, in Stem Cell Biology), in which this type of deviations have been alleged in the actions of the editorial committees involved. <sup>42</sup> In addition, deficient objectivity was confirmed in the analysis of allegations regarding breaches of confidentiality during the review process. <sup>3.4</sup>

The bias towards the successful result is another of the deficiencies attributed to PR, which greatly reduces the publication of negative or null results, <sup>43</sup> relevant for



the development of science. *As a solution, a model called result-blind PR* has recently been proposed, <sup>44</sup> which is based on the presentation of the manuscript to the reviewers in two stages: 1) the theoretical framework, the methodology and the format of tables and figures, without the results, and 2) with the real data, for evaluation once the first stage has expired. It should be implemented appropriately, in order not to unnecessarily lengthen the review process, and take advantage of the modular distribution of information in the future article. Due to its closeness in design, it could be integrated into the *Peer Agreement* model.<sup>1.25</sup>

A very particular case is the requests for additional experiments, which make the researcher waste time and resources without contributing new elements to the manuscript, unless they are experimental deficiencies. <sup>20</sup> These experiments should not comprise the next phase of the work, nor affect the conclusions or scope of the manuscript under evaluation, according to *Ploegh*. <sup>20</sup> However, the debate persists. For example, *Altschuler* posited that "collateral experiments [suggested by reviewers] are often better than those the authors plan to do in subsequent stages. They often reinforce the original results and lead to useful findings." <sup>Four. Five</sup>

#### QUALITY CONTROL OF PEER REVIEW

There are several approaches to evaluating the quality of PR, based primarily on survey studies. Although they have served to identify deficiencies in the PR, we must not lose sight of the fact that they are exploratory and do not constitute records of real cases. Hence, it is recommended to carry out real-time control at the level of editorial bodies of the deficiencies that arise, correctly documented and allowing effective measures to be taken to prevent them. <sup>46</sup>



*Gibson* et al <sup>47</sup> found during an author perception study that authors of rejected manuscripts provided less feedback to the editorial team about the quality of the PR. Therefore, any study of this type must avoid bias when considering opinions of authors of rejected works, which does not mean failing to include them. As demonstrated, the authors were able to discern between two types of information: the editorial processing and the reviewers' comments, the latter being the ones that raise the most awareness and, curiously, the most criticized by authors of rejected works.

There have also been deficiencies in studies evaluating errors during RP. *Jefferson* et al., in 2002, <sup>27</sup> found methodological inconsistencies in the studies reviewed up to that time that evaluated RP. In a subsequent analysis <sup>28</sup>, this same group found that PR demonstrably increased overall manuscript quality and readability in a limited number of studies. However, additional analyzes are required in order to make valid generalizations, without being limited by the deficiencies inherent to the PR model analyzed.

#### PR MANAGEMENT

It is valid to point out that editorial management between editors and authors is generally remote, although it greatly benefits from face-to-face interaction with the authors, especially in topics that are little covered or have recently appeared, and in the case of university journals, whose authors They are generally being trained in that role.  $^{48}$ 



Since the emergence of scientific journals, management was by postal correspondence, later replaced by digital computing and communications platforms as facilitators of all editorial processes. <sup>49</sup> Due to their wide scope, these have made it possible to estimate the interested public within the general PR process, with the proposal of the *Selected Articles Network* (SPN) being the most advanced example. <sup>fifty</sup>

One of those most notable pre-PR platforms is the arXiv repository, founded in 1991 by *Paul Ginsparg* and initially developed to deposit articles in mathematics, physics, and computer science, which subsequently incorporated other branches of science. <sup>51,52</sup> Its dynamics included the deposit of the works, which was used so that many were reviewed by the reading public prior to submitting them to PR, in a similar way to what occurs in the open process PR model (open *PR in two stages*<sup>6</sup> and SPN <sup>50</sup>). This influenced several journals in the subject fields related to the repository to modify their policies on pre-submission disclosure, which has been maintained regardless of the editorial management model they employ (pay-per-access or open access).

Computational developments led to today's online editorial management systems (also known as online PR systems, <sup>49,53</sup> for a historical approach to their development and a list of existing systems, respectively). The time saved by the decrease in the volume of documentary work favored requiring reviewers to spend more time training authors during the editorial interaction. <sup>49</sup>

The *Medical Journal of Australia* was the first attempt to harness the full potential of digital platforms and the Internet for PR, as early as 1996. <sup>54</sup> It was composed of *conventional PR* prior to online publication and *open-process* online PR. before



being published in print. This journal published the reviewers' reports along with the accepted articles on one page, and the authors added changes based on reader comments for the print publication. <sup>55</sup> Over time it led to the *Review Community* variant, of *open process RP*, with a secure forum. <sup>49</sup>

Today, in addition to online management systems, there are commercial editing and pre-RP services, <sup>56</sup> as well as automated formal manuscript review systems for authors and editors (e.g., PaperMaker), <sup>57</sup> that help provide better finishing prior to shipping. The possibility of delegating PR to professional flow management services (for example, Editorial Office Ltd. <sup>58</sup> and JournalPrep) has even been raised, <sup>59</sup> particularly in certain editorial situations.

### IMPLICATIONS OF DISCLOSING THE REVIEW

Although not absolutely, digital platforms have largely supplanted print publications and magazines as vehicles for disseminating debates about scientific veracity and the impact of science, <sup>60</sup> clearly relegating them as archival repositories of information. This effect is being counteracted by the growing incorporation of the most relevant approaches to the debates that occur on their digital platforms (forums) in the Comments and Opinions, <sup>44</sup> Letters to the Editor <sup>24</sup> and Correspondence <sup>45</sup> sections of scientific journals.

In another order, the citability of scientific articles has accelerated by being able to cite them independently of the journal through the use of persistent identifiers such as the DOI, <sup>61</sup> even since their disclosure to be evaluated by *open process* PR. <sup>6</sup> However, the latter implies care on the part of those who cite articles under review under the *open process* PR model, to always cite the published article and



not the manuscript under discussion (for example, *two-stage open PR*). <sup>6</sup> This model includes dissemination prior to its validation by the PR, but considers that it is a genuine scientific article when it is accepted after the PR and with the corresponding bibliographic data.

This new development forces us to discard the duplicate publication status in the event that a manuscript disclosed during *open process PR and resubmitted to a conventional PR* journal is rejected. It could also generate a subsequent evolution of bibliographic standards, to reflect the character of "disclosed in review", rejected or not, unless published later, something similar to what happens with publication online first, or with retraction . <sup>62</sup> This would add a greater level of complexity to the citation metric, although it could be effectively alleviated with digital identification codes (*Rodríguez EG*, manuscript in preparation). Another variant would be to remove the manuscript and its revisions from the journal's website, a practice established in the *Semantic Web Journal*, <sup>63</sup> that is, "unpublish" manuscripts rejected or delayed in responding to the open review after a period of eight weeks. , in conjunction with their reviews. Unpublication seems to be a more reasonable variant, as it avoids distortions in the citation metrics similar to those generated by retractions.

In general, PR is a process of quality assurance and control of the content of a journal, and serves as a consultation for new reviewers on certain patterns to consider to adequately evaluate experimental practices and results. At a more advanced stage of its implementation, the analysis of the information published on PR acts on the same topic could serve to compare different schools of science, taking into account the basic training and academic qualifications of the reviewers.



#### PEER REVIEW AS AN INDICATOR OF SCIENTIFIC PERFORMANCE

The act of PR remains largely anonymous and consumes a great deal of time. To solve this, variants have been suggested that measure the quality of the reviewer and publicly recognize their work, as part of their scientific curriculum. <sup>64</sup> It has been proposed that the change from *conventional PR* to *open process PR* would allow the reviewer's performance during the PR to be used as another parameter for evaluating their scientific career, by identifying themselves under their full name or digital identifier in the publication. <sup>14</sup> In this sense, it adds to the recent use of the Acknowledgments section for such purposes, in conjunction with authorship <sup>65</sup> and bibliometric indicators that qualify researchers as experts on a topic.

The use of reviewer performance as part of the scientific trajectory would only require disseminating information on the review process and the development of indicators that measure the parameters of the reviewer's function in the same journal and comparatively between journals (for example: number of reviews in a period, compliance with deadlines, methodological and formatting assessments, coverage of the sections of the manuscript, observations on statistical methods, tone of communication with the authors, use of literature to justify the observations). This would also favor the development of strategies to improve the portfolio of referees.

No less important, the study of the authors' responses also allows them to be evaluated in terms of their abilities to communicate adequately and face the



scientific debate with the reviewers, beyond their conviction or the contribution of the necessary technical elements.

A rating system for the work of reviewers was reported by *Kravitz* et al., <sup>66</sup> who classified the reviewers of the *Journal of General Internal Medicine* during the period from June 2010 to June 2011 into Reviewers and Senior Reviewers. This last category was determined by 2 or more reviews, a performance quality parameter never less than four on a quality scale of 1 to 6 as rated by the associate editors, and no more than 30 days for review delivery. In this line of thought, it has recently been proposed to establish a review efficiency index (or REF index) for each reviewer, <sup>67</sup> which is facilitated by the use of online editorial management systems.

Tools have also been developed to evaluate the performance of reviewers, associated with online editorial management systems. *In the case of open PR* models, there is the example of UCount, <sup>68</sup> a tool based on the Review Quality Instrument (RQI) <sup>69</sup> previously implemented by the editors of the *British Medical Journal*, and which is part of the editorial management system. online e-Scripts <sup>70</sup> (analyzed by *Birukou* et al.). <sup>71</sup> The system allows assigning values to the review act according to its relative quality on a 5-point scale in accordance with the parameters analyzed during the PR. However, the limitations of the RQI in assessing the comprehensiveness of the review have recently been recognized. <sup>72</sup>

The use of any reviewer performance indicator will depend on the consensus that exists in science systems about its need and relevance, since the validated communication of research results rests on the quality of the PR process. <sup>67</sup> This



would encourage reviewers to increasingly consider the invitation to review as a recognition of their expertise and not as an overload. <sup>16</sup>

### CONCLUSIONS

Unquestionably, PR is constantly refined to adapt to the changing requirements of scientific communities and the thematic evolution of research, without affecting its quality and purposes. Therefore, the study of editorial resubmission flows could facilitate a more precise vision of poorly characterized editorial dynamics, useful for predicting the relevance of the research result for readers once published, beyond citation-based indicators.

Likewise, the development of reviewer performance indicators will add value to the editorial work, outside the restricted framework of conventional review systems, thus adding to the authorship, citation analysis, and the Acknowledgments as endorsements section. of investigative performance.

The current evolution of PR points towards the management of the processing of scientific manuscripts and the validation of their data and metadata. In this context, scientific publishers are seen as centers of informational intelligence beyond the conventional or intrinsic evaluation of scientific texts. And the variants of pre-RP and *open process PR* in secure forums could reinforce the usefulness of unpublished manuscript repositories, by reducing their risks for the authors and depending on the scientific community in question.

#### Thanks



To the professors of the Diploma in Health Information Management in its 2011-2012 edition, belonging to the Virtual University of Health (Infomed, Cuba), for the knowledge provided on the use of the BIG6  $^{TM}$  model , from whose final evaluation exercise the present review article. To the Management and Information Services Directorate (GSI) of the Center for Genetic Engineering and Biotechnology (CIGB) of Havana, for facilitating participation in the diploma and technological support. To the reviewers and editors who evaluated the manuscript, for their anonymous but very accurate contribution.

### BIBLIOGRAPHICAL

### **REFERENCES** 1.

Rodríguez EG. Editorial peer review: Roles and processes [Internet]. RCICS. 2013[citedApril4,2013];24(2). Availableat: http://acimed.sld.cu/index.php/acimed/article/view/410

2. Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. JAMA. 2002;287(21):2805-8.

3. Fahy K. Perceived journal quality: An indicator of research quality. Women Birth. 2008;21(3):97-8.

4. Journal of Bacteriology [Internet]. 2013 [cited April 4, 2013]. Available at: <u>http://jb.asm.org/content/190/1/1.full</u>



5. Artificial DNA: PNA & XNA [Internet]. Landes Bioscience Journals. 2013 [cited April 4, 2013]. Available at: <u>http://www.landesbioscience.com/journals/artificialdna/guidelines/</u>

6. Sandewall E. Maintaining live discussion in two-stage open peer review. Front Comput Neurosci. 2012;6:9.

7. Burnham JC. The evolution of editorial peer review. JAMA. 1990;263(10):1323-9.

8. Ali J. Manuscript rejection: causes and remedies. J Young Pharm. 2010;2(1):3-6.

9. Lumb PD. Conflict of interest; disclosure; peer review. J Crit Care. 2011;26(4):333-4.

10. Akst J. I hate your paper. Scientist. 2010;24:36.

11. Calcagno V, Demoinet E, Gollner K, Guidi L, Ruths D, de Mazancourt C. Flows of Research Manuscripts Among Scientific Journals Reveal Hidden Submission Patterns. Science. 2012;338(6110):1065-9.

12. Scientific Reports [Internet]. 2013 [cited April 4, 2013]. Available at: <u>http://www.nature.com/srep/authors/index.html</u>

13. Sen CK. Rebound peer review: a viable recourse for aggrieved authors? Antioxid Redox Signal. 2012;16(4):293-6.

14.PulvererB.Peerreviews:somearealreadypublic. Nature. 2011;474(7351):285.

15. Saper CB. The Neuroscience Peer Review Consortium. Neurobiol Dis. 2009;33(3):313-4.

16. Khan K. Is open peer review the fairest system? No. BMJ. 2010;341:c6425.



17. Demystifying peer review. Nat Cell Biol. 2010;12(5):413.

18. NeurosciencePeerReviewConsortium. Informationforauthors[Internet]. 2013[citedApril4,2013]. Availableat: <a href="http://nprc.incf.org/authors/information-for-authors">http://nprc.incf.org/authors/information-for-authors</a>

19. Proceedings of the National Academy of Sciences of the United States of America. Editorial policies [Internet]. 2013 [cited April 4, 2013]. Available at: <u>http://www.pnas.org/site/authors/editorialpolicies.xhtml</u>

20. Ploegh H. End the wasteful tyranny of reviewer experiments. Nature. 2011;472:391.

21. Waltman L, van Eck NJ, van Leeuwen TN, Visser MS, van Raan AF. On the correlation between bibliometric indicators and peer review: reply to Opthof and Leydesdorff. Scientometrics. 2011;88(3):1017-22.

22. Jackson JL, Srinivasan M, Rea J, Fletcher KE, Kravitz RL. The validity of peer review in a general medicine journal. PLoS One. 2011;6(7):e22475.

23. Parker R, editor. Response to Sandewal's alternative to peer review. AAAS/UNESCO/ICSU Workshop on Developing Practices and Standards for Electronic Publishing in Science; Paris; 1998.

24. Young K. Rethinking peer review. CMAJ. 2010;182(3):281.

25. Loonen MP. Alternative peer review system: Peer agreement system. J Plast Reconstr Aesthet Surg. 2010;63(11):1931-2.

26. O'Dowd A. Peer review must stay as guarantee of quality, research leaders tell MPs. BMJ. 2011;342:d3656.

27. Jefferson T, Wager E, Davidoff F. Measuring the quality of editorial peer review. JAMA. 2002;287(21):2786-90.



28. Jefferson T, Rudin M, Brodney Folse S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. Cochrane Database of Systematic Reviews [Internet]. 2007(2); [cited April 4, 2013]. Available at: <u>http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/MR000016/f</u> rame.html

29. Schroter S, Tite L, Hutchings A, Black N. Differences in review quality and recommendations for publication among peer reviewers suggested by authors or by editors. JAMA. 2006;295(3):314-7.

30. Peering into review. Nat Med. 2010;16(3):239.

31. Kadar N. Systemic bias in peer review: suggested causes, potential remedies. J Laparoendosc Adv Surg Tech A. 2010;20(2):123-8.

32. Raff M, Johnson A, Walter P. Painful publishing. Science. 2008;321(5885):36.

33. Shattell MM, Chinn P, Thomas SP, Cowling WR. Authors' and Editors' Perspectives on Peer Review Quality in Three Scholarly Nursing Journals. J Nurs Scholarsh. 2010;42(1):58-65.

34. Resnik DB, Gutierrez-Ford C, Peddada S. Perceptions of ethical problems with scientific journal peer review: an exploratory study. Sci Eng Ethics. 2008;14(3):305-10.

35. Rangachari PK. Teaching undergraduates the process of peer review: learning by doing. Adv Physiol Educ. 2010;34(3):137-44.

36. Reynolds JA, Thompson RJ, Jr. Want to improve undergraduate thesis writing? Engage students and their faculty readers in scientific peer review. CBE Life Sci Educ. 2011;10(2):209-15.



37. Freda MC, Kearney MH, Baggs JG, Broome ME, Dougherty M. Peer reviewer training and editor support: results from an international survey of nursing peer reviewers. J Prof Nurs. 2009;25(2):101-8.