Peer review and journal impact factor: the two pillars of contemporary medical publishing

Triaridis S1, Kyrgidis A1,2
1 Department of Otolaryngology Head & Neck Surgery, Faculty of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
2 Department of Oral Maxillofacial Surgery, Faculty of Dentistry, Aristotle University of Thessaloniki, Thessaloniki, Greece

Abstract
The appraisal of scientific quality is a particularly difficult problem. Editorial boards resort to secondary criteria including crude publication counts, journal prestige, the reputation of authors and institutions, and estimated importance and relevance of the research field, making peer review a controversial rather than a rigorous process. On this background different methods for evaluating research may become required, including citation rates and journal impact factors (IF), which are thought to be more quantitative and objective indicators, directly related to published science. The aim of this review is to go into the two pillars of contemporary medical publishing, that is the peer review process and the IF. Qualified experts’ reviewing the publications appears to be the only way for the evaluation of medical publication quality. To improve and standardise the principles, procedures and criteria used in peer review evaluation is of great importance. Standardizing and improving training techniques for peer reviewers, would allow for the magnification of a journal’s impact factor. This may be a very important reason that impact factor and peer review need to be analyzed simultaneously. Improving a journal’s IF would be difficult without improving peer-review efficiency. Peer-reviewers need to understand the fundamental principles of contemporary medical publishing, that is peer-review and impact factors. The current supplement of the Hippokratia for supporting its seminar for reviewers will help to fulfil some of these scopes. Hippokratia 2010; 14 (Suppl 1): 5-12

Key words: impact factor, peer-review, citation, editor, medical, quality

Corresponding author: Stefanos Triaridis, 215 Lambraki St, Thessaloniki, 543 52, Greece, Tel. +30-6936949394, E-mail: triaridis@hotmail.com

The appraisal of scientific quality of medical publications is a particularly difficult problem, as published scientific results ought to be scrutinised by experts in the field and qualitative and quantitative scores to be provided according to well-established rules. In real life, editorial boards usually perform what is called peer review with general competence rather than with the expert’s view that is needed to assess primary research data1-2. Editorial boards resort to secondary criteria including crude publication counts, journal prestige, the reputation of authors and institutions, and estimated importance and relevance of the research field, making peer review a controversial rather than a rigorous process3-6. On this background different methods for evaluating research may become required, including citation rates and journal impact factors, which are thought to be more quantitative and objective indicators, directly related to published science2. The citation data are obtained from a database produced by the Institute for Scientific Information (ISI) in Philadelphia, which continuously records scientific citations as represented by the reference lists of articles from a large number of the world’s scientific journals since 19717. These reference lists are reorganized to illustrate how many times each publication has been cited within a certain period, and by whom, and the results are published as the Science Citation Index (SCI)8. Based on the Science Citation Index and authors’ publication lists, the annual citation rate of papers by a scientific author or research group can thus be calculated9. Similarly, the citation rate of a scientific journal-known as the journal impact factor-can be calculated as the mean citation rate of all the articles contained in the journal10,11. Journal impact factors, which are published annually in SCI Journal Citation Reports10, are widely regarded as a quality ranking for journals and used extensively by leading journals in their advertising11,12. In this manuscript, which is included in a special supplement aimed to possible peer-reviewers of our journal, along with a series of papers providing guidance related to academic skills for our readers, we aim to go into the two pillars of contemporary medical publishing, that is the peer review process and the journal impact factors.

I. Peer Review

History of peer-review

The maturation of peer review process was slow and somewhat chaotic13. Different editors engaged various peer review techniques. For instance, the Lancet did not utilize any peer review process prior to 1976 as the editors considered it unimportant. The Journal of the American
Medical Association (JAMA) processed their submissions through an in-house review board and only on rare occasions would they send manuscripts to outside experts. The British Medical Journal (BMJ) pioneered the peer-review process by sending every non-editorial submission to an external recognized expert from 1893. By the late 20th century, peer review was widely implemented and is currently adopted by the majority of biomedical journals. Over this period, many journals were obliged to adopt the peer review process by the increasing specialization within research fields and the increasing competition among journals for quality manuscript submissions. Nowadays, most biomedical researchers consider peer review a necessity for research articles. Contemporary peer review systems have evolved from their 18th century counterparts, and there are considerable variations among them. Peer-review systems universally include a process of systematically distributing, evaluating, and reaching a consensus on the qualities of submitted manuscripts, leading to publication following acceptance (with or without revision), or rejection. The foundations of this process are the editors and the reviewers. Most reviewers acquire their training in manuscript reviewing not through any kind of formal educational process but simply by undertaking it. Medical research training rarely incorporates instructions in the art of peer reviewing. Editors usually assign reviewers based on their expertise in a given field and their availability. If a reviewer repeatedly submits incompetent reviews, it is unlikely that he will continue to receive invitations for peer review of manuscripts.

The ability of reviewing manuscripts is one that improves with practice, similarly to any kind of human dexterity. While peer reviewing is not based on subject’s talent, certain characteristics such as evenness, meticulousness and honesty contribute to this dexterity. The techniques of peer reviewing can be taught and put into practice. Peer review is a recognized and critical component of the overall publication process that confers “added value” to a submitted manuscript. Table 1 proposes some objective criteria aiming to simplify to a certain degree the task of peer-reviewers.

### Table 1: Checklist for peer-reviewers.

<table>
<thead>
<tr>
<th>Section</th>
<th>Specific Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title, Authors</strong></td>
<td></td>
</tr>
<tr>
<td>Title is:</td>
<td></td>
</tr>
<tr>
<td>• clear and informative;</td>
<td></td>
</tr>
<tr>
<td>• representative of the content and breadth of the study;</td>
<td></td>
</tr>
<tr>
<td>• captures the importance of the study and the attention of the reader.</td>
<td></td>
</tr>
<tr>
<td>Number of authors appears to be appropriate.</td>
<td></td>
</tr>
<tr>
<td><strong>Abstract and Keywords</strong></td>
<td></td>
</tr>
<tr>
<td>Abstract is complete. Results in the abstract are presented in sufficient detail. The conclusions in the abstract are justified by the information in the abstract and the text. There are no inconsistencies in detail between the abstract and the text. All of the information in the abstract is present in the text.</td>
<td></td>
</tr>
<tr>
<td>Keywords are appropriate and sufficient.</td>
<td></td>
</tr>
</tbody>
</table>

**Accepting a review**

Editors depend on the assistance of competent reviewers to whom they rely on for manuscript quality control. Reviewers are motivated by a sense of duty, altruism, and a desire to contribute in an important way to the safeguarding of high standards and truth in their specific research fields. Being invited to review a manuscript is an honour, not only because one is being recognized as an expert in a particular research field but also because of the responsibility and service he provides to the journal and scientific community.

A reviewer needs to act as a defender of the submitting author, rather than an opponent. This acting is best summarized by Benos et al in the phrase “a reviewer should treat a manuscript being reviewed as he/she would want his/her own paper treated”. At the same time the reviewer must also act as the “journal’s advocate” and make sure that the best possible research is published.

The purpose of peer review is to ensure: i) quality, checking that no mistakes in procedure or logic have been made; ii) that the results presented by authors support the conclusions drawn; iii) that no errors in citations to previous work have been made; iv) that all human and animal protocols were conducted according to good clinical practise and following approval of institutional review boards; and, very importantly, 5) that the work is original and significant.

If a reviewer disagrees with the conclusion of an author, it is his/her obligation to provide definitive reasons or appropriate citations, rather than simply make remarks of disbelief of author’s data. If a reviewer is biased against the author, he should refrain from reviewing the paper. A reviewer must know the topic and understand the context in which the study was done. As many manuscripts nowadays are collaborative efforts between different laboratories using lots of different techniques, it is questionable that any single reviewer will be expert in all of the protocols encountered in a particular manuscript. The reviewer should comment only on those aspects of the scientific work with which he is familiar, at the same time informing the editor about this. For example, the reviewer may provide a series of comments of the research field and further prompt the editor to assign the manuscript to a third reviewer with statistical knowledge."
**Introduction**

The introductory literature review is up-to-date, references are mainly primary sources and a logical context is defined. The last paragraph of the introduction clearly states the aim of the study and the research hypothesis.

**Methods**

**Design**

The design is defined and clearly described, and is sufficiently detailed to permit the reproduction of the study for the research question.

**Instrumentation, Data Collection**

The measurement of instruments is appropriate given the study’s variables.

**Population and Sample Size**

The population is clearly defined; sampling procedures are sufficiently described to allow for reproducibility. Subject and experimental samples are appropriate to the research question. Various types of bias are addressed.

**Data Analysis and Statistics**

Data analysis procedures:
- are sufficiently described and are sufficiently detailed to permit the study to be reproduced;
- conform to the research design; hypotheses, models, or theory drives the data analyses.

The assumptions underlying the use of statistics are fulfilled by the data, such as measurement properties of the data and normality assumptions. Statistical tests are appropriate. If statistical analysis involves multiple tests or comparisons, proper adjustment of significance level for chance outcomes was applied. Power issues are considered in statistical studies with small sample sizes.

**Results**

**Presentation**

Results are:
- organized in a way that is easy to understand;
- presented effectively;
- are complete, accurate (e.g., numbers add up) sufficient and appropriate;
- tables, graphs, or figures are used judiciously and agree with the text;
- tables contain new information and are parsimonious;
- figures are original, of high quality.

**Reporting of Statistical Analyses**

The assumptions underlying the use of statistics are considered given the collected data. The statistics are reported correctly and appropriately. The number of analyses is appropriate for the sample size.

**Discussion**

The discussion:
- topics are clearly stated; key points stand out;
- follows from the design, methods, and results;
- justification of conclusions is well articulated;
- interpretations of the results are appropriate;
- the conclusions are accurate;
- the study limitations are discussed;
- Statistical differences are distinguished from meaningful differences;
- Practical significance or theoretical implications are discussed.

**References**

Reference citations are complete and accurate, according to journal style.

**Generic Criteria**

**Relevance**

The study:
- is relevant to the mission of the journal or its audience;
- addresses important issues;
- adds to the literature already available on the subject;
- and has adequate generalizability because of the selection of subjects, setting, and intervention.

**Presentation and Documentation**

The text is:
- well written and easy to follow.
- The vocabulary is appropriate.
- The content is complete and congruent.
- The manuscript is well organized.

**Scientific Conduct**

- There are no instances of plagiarism.
- Ideas and materials of others are correctly attributed.
- Prior publication by the authors of substantial portions of the data or study is appropriately acknowledged.
- There is no apparent conflict of interest.
- There is an explicit statement of approval by an institutional review board for studies directly involving human subjects or data about them.
knowledge. Table 2 summarizes the responsibilities of a peer-reviewer, which the reviewer needs to fully understand prior to accepting a review. Depending on the type of the article (original report, review, case report, letter etc.), the review process may require 1–3 hours and 500 to 1,000 words.

Sometimes reviewers base their judgments on cues that have only a weak relation to quality such as statistical significance, large sample size, complex procedures, so-called “negative” data, and obscure writing. The publication by Atkinson et al who prepared three versions of a fake manuscript in which identical findings differed only by the degree of statistical significance and assigned them to peer-reviewers showed that the reviewers recommended rejection of the paper with non-significant findings three times as often as those with significant findings. The latter experiment confirmed that reviewers place too much emphasis on statistical significance.

On the other hand, statistical significance probably means results that would allow for the reproduction of the experiment and thus advancement of current knowledge. Manuscripts with statistically significant results tend to be cited more frequently, but this also leads to the so-called “publication bias” where studies with statistically “non-significant” results find it harder to get published.

### Declining a review

The first thing a reviewer must do when accepting an invitation to review a manuscript is to agree or decline. There may be several reasons to decline the invitation to review. The reviewer may have conflict of interest that would prevent him from writing truly fair and objective review. Further, the topic or methodology may fall outside of the reviewer’s knowledge and therefore ability to adequately appraise it. Finally, it may not be possible for the reviewer, given other time constraints or commitments, to complete the review within the expected time frame. At this point it must be emphasized that the sooner the review is performed, the better for both the author and the journal. Quick peer review process expedites scientific knowledge spread and allows for a higher journal impact factor; thus it needs to be widely endorsed. In the case reviewers decline they can often recommend other experts for the review. Manuscripts released to reviewers are confidential, and their contents should never be discussed with other colleagues until they are published.

#### Decision on publication

A common mistake often made by reviewers is to assume that the manuscript will be rejected and therefore to provide little feedback to the authors. Regardless of their recommendation, reviewers are anticipated to provide constructive feedback that will aid authors in the future submission of their work to another journal. Ultimately, the decision concerning publication belongs to the editors who often receive ambiguous and contradictory reviews. Therefore, reviewers should never predict the editorial decision in their “comments to the authors”.

#### Blinding reviewers or not?

Whether reviewers need to be blinded and if this is feasible is another matter of concern. The publication bias favouring prominent researchers from well-reputed institutions has been confirmed by the much-cited study by Ceci and Peters, which blinded reviewers to published papers from well-known research groups and found high rates of manuscript rejection based on scientific grounds. However, other studies reported that while there was an association between high institutional status and acceptance of brief reports, this relationship was not observed in the acceptance rates of regular articles. On the contrary, the likelihood of recommendation for acceptance and of selection for publication of brief reports was found to correlate with the prestige of the institution.

Regarding the so-called gender bias, it has been reported that female-authored manuscripts were accepted...
significantly more often by female reviewers rather than by male reviewers, while male reviewers did not discriminate between male- or female-authored manuscripts in terms of acceptance rate. Female editors had larger workloads and were more likely to reject manuscripts summarily.

Even when blinding exists, it is not a matter-of-course that the reviewer will not be aware of the author’s identity. Self-referential writing and small research fields are the most commonly given reasons for unsuccessful blinding.

**Truth**

Peer review cannot necessarily ensure that a paper is truthful. It can only claim that it is worth publishing. On the other hand, an effective peer-review process weeds out substandard work and methodological errors and blunts possible biases by scientific investigators.

**Review articles utility**

The expansion of the scientific literature has produced a concomitant increase in the number of review articles. Review articles are heavily cited and inflate the impact factor of journals. One may posit that the sheer number of review articles belies their function. The reality that many review articles are poorly cited has been reported to raise concerns about the harm that poor review articles can cause. First by making it harder to discriminate good reviews and secondly by propagating scientific error through lack of critical appraisal of original research.

A review article should not provide just a rehash of the literature. Rather, the author of a review shapes the literature of a field into a story in order to enlist the support of readers to continue that story. The review selects from a wide list of research papers, puts them side by side and creates a story that combines them, similar to a theatre play with actors and events but still without an ending. Thus, it attracts the reader into the writer’s point of view of what has happened, and by arranging the current information, suggests what can be done in the future. Therefore, reviews in human sciences do not merely summarize, but they also allow the construction or reconstruction of knowledge claims.

**Reviewers are authors, authors are reviewers**

Most authors utilize the reviewer’s comments to modify their manuscripts before submitting them to another journal. This is beneficial because it can alleviate the authors’ feelings for the peer-review process following a rejection. Furthermore, given that not all articles are published and that each submitted article is usually reviewed by at least two reviewers, any expert author is expected to also perform reviews, at least half times of his published papers. Thus, reviewers and authors are not two different groups. Rather, they belong to the same group serving one role or the other, for the good of scientific knowledge. Thus, both authors and reviewers need to be courteous and treat others fairly. It is the only way that they will not find themselves in a disagreeable situation, since role exchange is a necessity.

**Online submission**

Currently many scientific journals receive and process manuscripts online. It has been reported that online manuscript submission is more efficient than hard copy submission for all manuscript types.

**II. Impact factor**

**Definitions**

A major indicator of submitting a paper for publication is the journal impact factor (IF) issued yearly by the Institute for Scientific Information (ISI) and the journal ranking in the relevant subject category of the Journal Citation Reports. In plain language, the impact factor represents the average number of citations per article the journal received during the previous 2 years. The impact factor is only one of three standardized measures created by the ISI which can be used to measure the way a journal receives citations to its articles over time (the others are the immediacy index and the cited half-life). The build-up of a journal’s citations tends to follow a skewed-to-the-left curve like the one presented in Figure 1.

![Figure 1: Graphical presentation of the three indicators by the Institute for Scientific Information (ISI).](image)

Citations to articles published in a given year often peak between two and six years after publication. After this peak citations tend to decline over time. The citation curve of any journal can be described by the relative size of the curve (area under the curve), the extent to which the peak of the curve is close to the origin, and the rate of decline of the curve. These characteristics form the basis of the ISI indicators: impact factor, immediacy index and cited half-life. The immediacy index gives a measure of the skewness of the curve, that is, the extent to which the peak of the curve lies near to the origin of the graph. It is a fraction where the citations a journal receives in the current year is the nominator and the number of articles it publishes in that year is the denominator. The cited half-
life is a measure of how long articles in a journal continue to be cited after publication. It is the number of years that the number of current citations takes to decline to 50\% of its initial value (the cited half-life is 6 years in the example given in Figure 1)

### Flaws

Impact factors are only one of a number of measures available for describing the “impact” that particular journals can have in the research literature. The value of the impact factor is associated with the subject area, type and size of a journal\textsuperscript{11}. Journal impact factors are not statistically representative of individual journal articles. Rather, journal impact factors correlate poorly with actual citations of individual articles\textsuperscript{2} meaning that the citation rate of a given article determines the journal impact, but not vice versa.

The IF, as an absolute figure is not always a reliable indicator as to which journal has actually scored better than the others\textsuperscript{12}. Some measures that have been proposed to better evaluate the IF of a particular journal include the fraction of self-citations of previous 2 years divided by the total-citations of all years, the self citations of the past two years divided by the number of citations in all years of journal history, the most citing journal other than same journal and the number of different citing journals\textsuperscript{12}. Calculating some or all of these, authors who are keen on statistics may be able to “predict” which would be the best journal to publish their work. Generally, authors searching for a journal to publish their articles should look beyond the absolute figures of the given impact factors and the rankings of the journals, which are often influenced by a variety of unclear features. In judging an individual journal, what counts should be its actual merits, not its impact factor\textsuperscript{12}. On the other hand, the fact that authors use criteria other than IF when submitting to journals further questions the accuracy of IF as a measure of quality, since authors of well-designed studies, may elect to publish in lower IF journals\textsuperscript{27-28}.

With regard to the calculation of the IF, only those classified as ‘articles’ or ‘reviews’ and ‘proceedings papers’ are counted in the denominator, whereas citations to all papers (including editorials, news items, letters to the editor, etc) are counted for the nominator\textsuperscript{11}. However, citations to “non-citable” items are still included in the database, meaning that self citations are not corrected for\textsuperscript{28}. Coverage of the database is not complete. For instance, books are not included in the database as a source for citations. Furthermore the database has an English language bias and is dominated by American publications\textsuperscript{2}. A clear correlation between the extent of the impact factor fluctuation between years and the size of the journal has been reported\textsuperscript{11}. Short publication lag has been reported to allow many short term journal self citations and give a high journal impact factor\textsuperscript{2}.

Impact factor is a function of the number of references per article in the research field\textsuperscript{2} meaning that IF figures are comparable only within the same field. Basic life research journals have the highest mean impact factor, followed by clinical medicine, pharmacology, and biological sciences\textsuperscript{11}. In general, surgical specialties journals, have lower impact factors than medicinal specialties journals. The depth of a research field relies for the most part to the commercial interest for new diagnostic and treatment modalities. Thus, research fields with literature that rapidly becomes obsolete are favoured\textsuperscript{2}. In example, oncology specific journals are favoured while small research fields tend to lack journals with high impact. Relations between publishing fields (for example, clinical versus basic research) strongly determine the journal impact factor. Since the impact factor depends on the expansion or the contraction of a given research field, comparative analyses among different research fields should therefore be adjusted to take into consideration the citation density (mean number of citations per article), half-life (number of years, prior to the current, covering half of the citations given to the journal during the current year), and special citation practices that are specific and unique to each scientific field\textsuperscript{12}.

Editors wish to achieve the highest possible impact factors for their journals. A high impact factor is –at first glance- erroneously taken as a measure of publicity, meaning higher financial income for the journal to expand\textsuperscript{27-28}. Various techniques have been proposed for IF manipulation: i) requiring revision of the manuscript references section and inclusion of articles published in the editor’s journal or affiliate journals (Selective journal self citation bias)\textsuperscript{27-28}; ii) publishing summaries of articles with relevant citations to them\textsuperscript{27}; iii) inflating self-citation through editorials and readers’ comments on published articles\textsuperscript{27}; iv) publishing articles that add citations to the nominator but which are not counted as “citable”\textsuperscript{27,28}; v) publishing a larger percentage of review articles over less-cited articles, including original research and, especially, case reports\textsuperscript{2,11,24,28}; vi) rejecting studies with non-significant results, regardless of their quality (publication bias)\textsuperscript{12,27-28}; vii) favouring the acceptance of articles originating from large and scientifically active research groups as well as articles with a large number of authors\textsuperscript{27,28}; viii) attracting the work of renowned scientists and leaders of research regardless of the real quality\textsuperscript{27,28}; x) publishing mainly popular science articles that deal with modern topics\textsuperscript{27,28}.

### Discussion

The process of peer review has undergone many changes since its introduction in the 18\textsuperscript{th} century. In the land of innovation and scientific research, traditional models continuously evolve. The current system of peer review is probably not perfect, while most scientists agree it is necessary. Editors of the journals and many scientists consider the system’s expense and time consumption worthwhile in the belief that it weeds out poorly designed studies and methodological errors and blunts possible biases by researchers\textsuperscript{1,4,14}. However, the core assumptions innate in the peer-review process must
be continuously evaluated and adapted to the changing environment. Although peer review is often viewed as the gatekeeper to the land of certainty, an editorial board and reviewers may not be able to make that distinction in a few months.

There are means for providing good constructive reviews. Some of these are best taught through mentors and reviewing practice. However, the most important traits of a potential reviewer are courtesy, evenness, and punctuality. Thus, peer reviewers, need to follow the golden rule: treat other manuscripts, as they would want their own to be treated. The entire peer review process, which in essence determines the public record of science, is based on trust; trust between authors and editors and trust between editors and reviewers. The quality and integrity of the entire scientific publishing enterprise depends in large measure on the quality and integrity of the reviewers.

Reviewing is both a privilege and responsibility. Preparing a useful, critical review may be time-consuming. Moreover, it clearly is a service to the journal, to the authors, to science at large, and to the reviewer because the reviewer becomes privileged to reading the latest in cutting-edge research. It has been suggested that, regardless of the perceived pre-eminence of any particular journal, reviewers should approach the review of each research paper the same way. This would allow for the improvement of journals overall quality and -hopefully- impact factor.

Editors and publishers are expectedly concerned about the quality, prestige, and even commercial success of their scientific journals. Furthermore, the IF is perhaps the most valuable index used to date that measures and quantifies journals’ quality and serves at the same time as a marketing tool. The use of absolute values of impact factors, outside of the context of other journals within the same subject area, is virtually meaningless; journals ranked top in one field may be bottom in another. To project the use of the journal impact factor from the journal to the authors of certain papers in the journal, implicates a lot of assumptions reducing its value. Medical journals contain many more types of source items (case reports, letter to the editor, quiz) than the standard scientific journal. Impact factors are useful in establishing the influence journals have within the literature of a certain discipline. Even so, they are an indirect measure of quality and should therefore be used with care.

For evaluation of scientific quality, there seems to be no alternative to qualified experts reviewing the publications. To improve and standardize the principles, procedures, and criteria used in peer review evaluation is apparently of great importance and the scientific community would be well served if efforts could be concentrated on this, rather than on developing ever more sophisticated versions of “impact” indicators. Standardizing and improving training techniques for peer reviewers, would allow for the magnification of a journal’s impact factor. This may be a very important reason that impact factor and peer review need to be analyzed simultaneously. Improving a journal’s IF would be difficult without improving peer-review efficiency. Peer-reviewers need to understand the fundamental principles of contemporary medical publishing, that is peer-review and impact factors.

Based on the above, the best possible route to increase a journal’s reputation and impact factor would be to train the available reviewers and potential future authors. For that, the role of medical journal clubs is essential. Medical journal clubs have been in place for over 100 years. Deenadayalan et al reported that journal club intervention was effective in improving knowledge and critical appraisal skills of potential reviewers. Characteristics of successful journal clubs include regular and anticipated training meetings, mandatory attendance, appropriate meeting timing and incentives, a trained journal club leader to choose papers and lead discussion, circulating papers prior to the meeting, using the internet for wider dissemination and data storage. The current supplement of the Hippokratia for supporting its seminar for reviewers will help to fulfill some of these scopes.

References

2. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ. 1997; 314: 497.