# PARTICLES & HEALTH 2021

## The Science of Scientific Publishing:

## Metrics, Misuses, Modifications and More

*Paul Brandt-Rauf, Sc.D., M.D., Dr.P.H*

*Distinguished University Professor and Dean of Biomedical Engineering Science and Health Systems*

*Drexel University*

Paris
October 20, 2021

# Metrics – In the beginning there was the impact factor

- 1955 – Eugene Garfield publishes a paper suggesting a grading system for journals based on the impact factor, the number of citations an average article in each journal receives
- 1960 – Garfield and a colleague start a company called ISI to tabulate citation statistics and calculate journal impact factors
- 1992 – ISI is bought by the Thomson Corporation and the impact factor is highly publicized

# Metrics - Calculation of the impact factor

IF for Journal X for Year A = (Citations in Year A to Articles Published in X in Years A-1 and A-2)/ (Articles Published in X in Years A-1 and A-2)

IFs based on >27 million citations for 5968 physical science journals and 1712 social science journals

Ifs range from <0.1 to >50 with half of journals <1

# Metrics -Top journals by impact factor

| | |
|---|---|
| Annual Review of Immunology | 52.4 |
| Cancer Journal for Clinicians | 44.5 |
| NEJM | 38.6 |
| Nature Reviews Cancer | 36.6 |
| Physiological Reviews | 33.9 |
| Nature Reviews Molecular Cell Biology | 33.1 |
| Reviews of Modern Physics | 32.8 |
| Nature Reviews Immunology | 32.7 |
| Nature | 32.2 |
| Science | 31.9 |
| Annual Review of Biochemistry | 31.5 |
| Nature Medicine | 31.2 |
| Cell | 28.4 |
| Nature Immunology | 27.6 |
| JAMA | 24.8 |

# Metrics - The value of the impact factor

- Higher ranking journals do get the message out better – in 2001 Matthew Stanbrook of the University of Toronto tracked what happened when 12 medical journals published a joint statement on research authorship and sponsorship, and, over the next 26 months, the highest IF journal received 100x as many citations as the lowest one for this identical statement
- So it can have a use for scientists, editors, publishers and others in judging journals

# Metrics - Evolution of the use of the impact factor

Evaluation of individual papers or researchers

In England, hiring panels routinely consider the IF of journals applicants publish in

In Spain, by law researchers are rewarded for publishing in journals in the upper third of Ifs

In China, scientists get cash bonuses for publishing in high IF journals

# Metrics - Evolution of the use of the impact factor

In China, graduate students in physics must place at least 2 articles in journals with a combined IF of 4 to get their PhDs

In the US, junior faculty are being told that to get tenure they should publish in high IF journals

# Misuses - Problems with the impact factor: effect on researchers

Is it fair to evaluate individual papers or researchers by IFs?

The IF applies to all papers in a journal over a year, not to a single paper, let alone to any author, and the distribution is highly skewed

The effects may be worse in certain areas that we are trying to promote, such as interdisciplinary research

# Misuses - Problems with the impact factor: effect on publishing

Do IFs adversely affect publication decisions?

IFs influence what gets published:

C. DeAngelis, editor of JAMA, says "editors of some top journals…won't publish articles because it won't help their impact factor"

F. Godlee, editor of BMJ, says "editors may be rejecting not only studies in smaller or less fashionable fields, but also important papers from certain regions of the world out of fear that such reports won't attract sufficient citation attention."

# Misuses - Problems with the impact factor: effect on publishing

Do IFs adversely affect publication decisions?

Editors "game" the system to increase IFs:

Publish more review articles at the expense of original research articles

Publish more uncounted items at the expense of original research articles

Do more press releases

# Misuses - Problems with the impact factor: effect on publishing

Do IFs adversely affect publication decisions?

Promote self-citation

- Editorials that cite numerous articles from previous issues give a noticeable jump in IF
- Editors may unethically pressure authors to cite articles from their journal to increase IF

# Misuses - Problems with the impact factor: effect on science

Do IFs adversely affect the course of scientific research?

IFs influence choice of research direction skewing the course of research:

Top journals require that papers be topical in addition to presenting important science so researchers shift the kinds of questions they investigate

The system slows the pace of science so less research gets done

# Modifications – Suggested improvements

- Change the system

ISI could count citations only to original research articles eliminating the problem of reviews, news stories, editorials and other kinds of material

ISI could lengthen the period covered to accommodate slower moving fields (e.g., IF has recently included a 5 year period option)

# Modifications - Suggested improvements

- Educate about the uses/abuses of IFs

Over a long term and in the aggregate it is one indicator of how well a journal disseminates information but only within the framework of a single scientific discipline

It should not be taken as an unequivocal measure of the scientific quality of individual articles in a journal

It should not be used for graded evaluations of individual scientists in terms of hiring, promotions or obtaining grants

# Modifications – A better journal citation metric: the Eigenfactor

- The Eigenfactor assumes the influence of a journal is best measured by the number of independent citations it attracts from other influential journals over an extended period

Citations a journal receives from other journals are weighted by the importance of the citing journal using the Eigenvector centrality that is calculated recursively such that values are transferred from one journal to another in the network until a steady-state equilibrium is reached (like Google's PageRank)

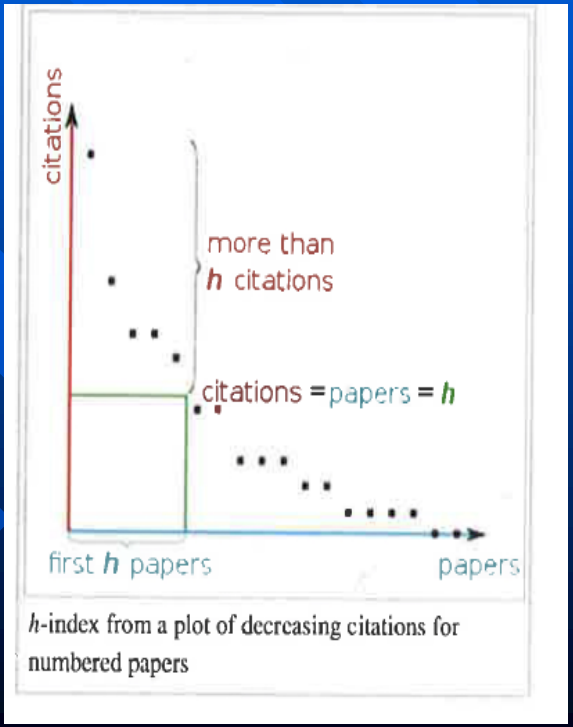
Eigenfactor is based on citations made in a given year to papers published in the prior five years reducing year-to-year volatility

Self-citation is eliminated disincentivizing bad referencing behavior

# Modifications – A better author citation metric: the h-index

- The h-index (Hirsch-index) is an author-level metric that attempts to measure both the productivity and impact of the publications of a scientist first proposed in 2005 by Jorge Hirsch at UCSD

  A scientist with an index of *h* has published *h* papers each of which has been cited in other papers at least *h* times, thus reflecting both number of publications and the number of citations per publication; it works properly only for comparing scientists working in the same field. For example, if a scientist has 5 publications with 10, 8, 5, 4 and 3 citations, respectively, the *h* index is 4 because the 4th publication has 4 citations and the 5th only has 3.

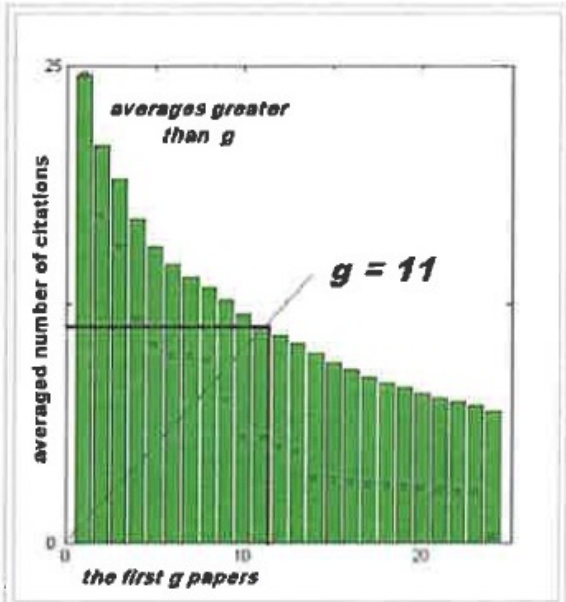


*h*-index from a plot of decreasing citations for numbered papers

# Modifications – A better author citation metric: the i10-index and g-index

- The i10-index is a simpler metric that just refers to the number of papers with 10 or more citations
- The g-index is a more complicated metric proposed by Leo Egghe in 2006

  It is based on the distribution of citations received by a scientist's publications, such that given a set of articles ranked in decreasing order of the number of citations that they received, the index is the unique largest number that the top $g$ articles received together at least $g^2$ citations.




An example of a g-index (the raw citation data, plotted with stars, allows the h-index to also be extracted for comparison).

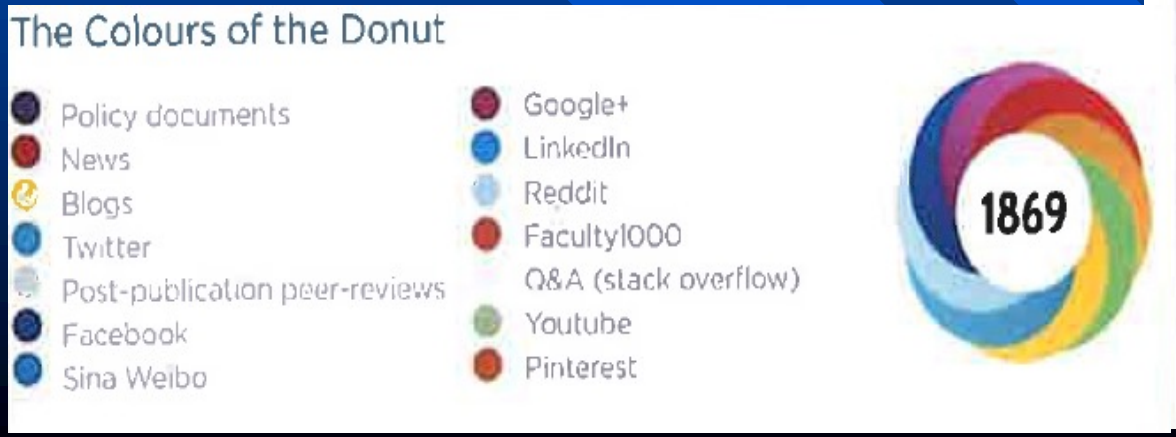
# Modifications – A better author citation metric: the i10-index and g-index

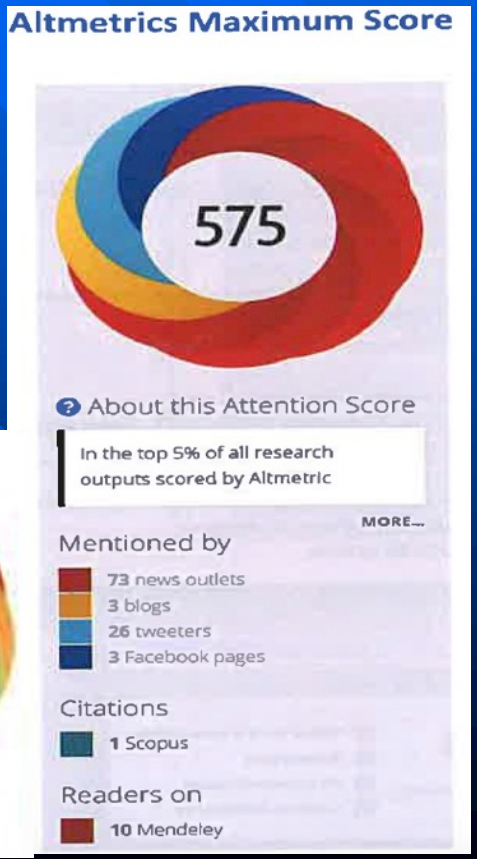
- The g-index is equivalently defined as the largest number *n* of highly cited articles for which the average number of citations is at least *n*. In that sense it differs from the h-index which doesn't average the number of citations; the h-index only requires a minimum of *n* citations for the least-cited article in the set and thus ignores the citation count of very highly cited papers. Roughly, the effect is that h is the number of papers of a quality threshold that rises as h rises; g allows citations from higher-cited papers to be used to bolster lower-cited papers in meeting this threshold. Therefore, in all cases g is at least h and in most cases higher. However, unlike the h-index, the g-index saturates whenever the average number of citations for all published papers exceeds the total number of published papers.

# Modifications – Measuring impact beyond citations: altmetrics

- Altmetrics was proposed in 2010 as a non-traditional alternative to citation-based metrics by tracking the attention that research outputs such as scholarly articles receive online

It includes other aspects of impact of a work such as how many data and knowledge bases refer to it, article views, downloads or mentions in social media and news outlets to derive a weighted score.



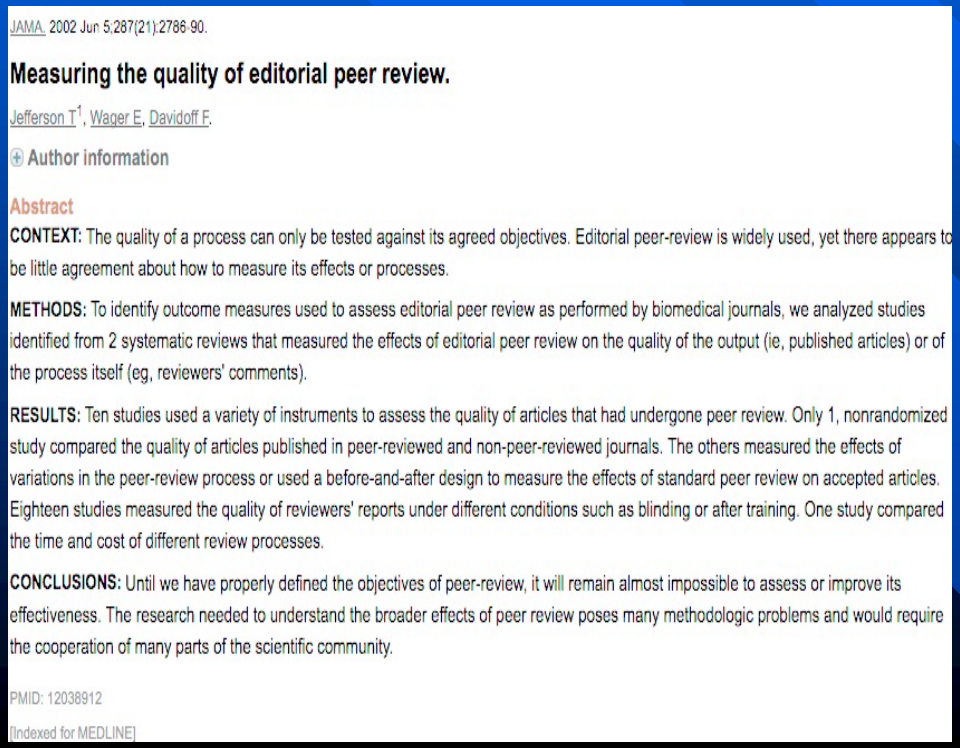
# Modifications – Measuring impact beyond citations: altmetrics

- Altmetrics are more difficult to standardize than citations. Altmetrics can be gamed since likes and mentions can be bought and are prone to self-citation. Altmetrics do not tell you anything directly about quality (a slight correlation has been found between mentions and likes and citations in the scientific literature; and number of tweets in the first 7 days of publications was a good indicator of highly cited articles)

- Nevertheless, funders, including the UK Medical Research Council, has shown interest in using altmetrics. Some universities, including the University of Pittsburgh, are experimenting with altmetrics for faculty promotion review.

# More – The future of scientific publishing?

- A cornerstone of quality control in scientific publishing has been the peer-review process, but the process has come under increasing scrutiny and criticism.



JAMA. 2002 Jun 5;287(21):2786-90.

**Measuring the quality of editorial peer review.**

Jefferson T[1], Wager E, Davidoff F.

Author information

Abstract

**CONTEXT:** The quality of a process can only be tested against its agreed objectives. Editorial peer-review is widely used, yet there appears to be little agreement about how to measure its effects or processes.

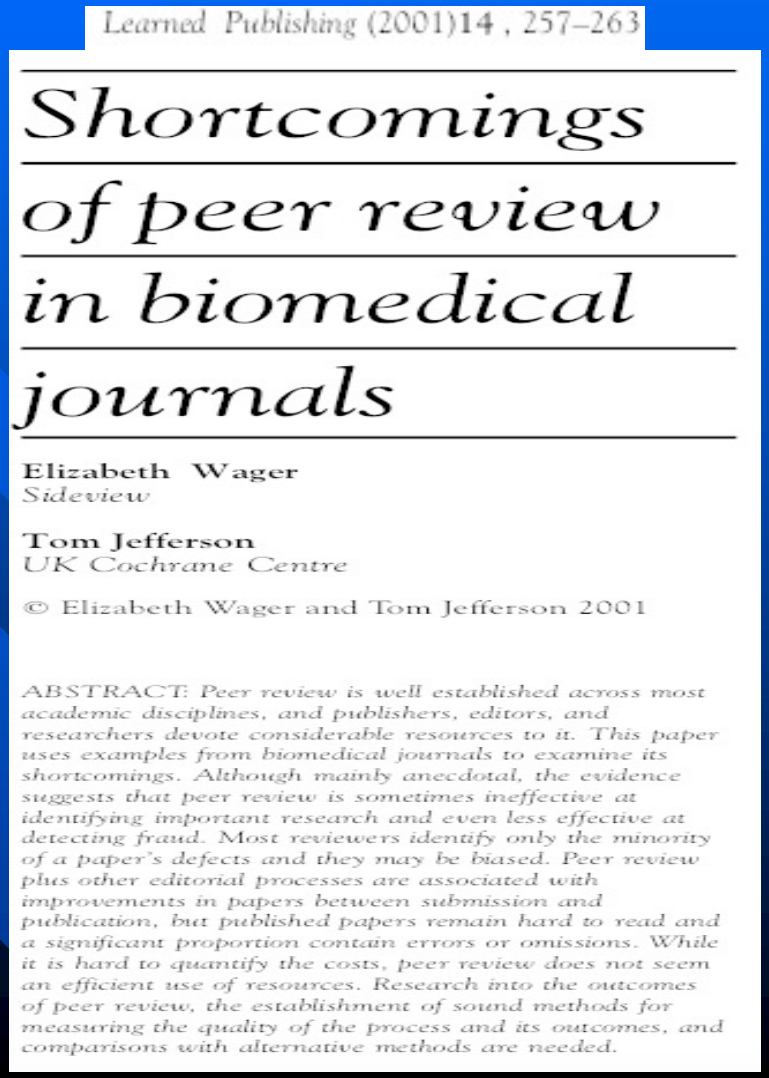
**METHODS:** To identify outcome measures used to assess editorial peer review as performed by biomedical journals, we analyzed studies identified from 2 systematic reviews that measured the effects of editorial peer review on the quality of the output (ie, published articles) or of the process itself (eg, reviewers' comments).

**RESULTS:** Ten studies used a variety of instruments to assess the quality of articles that had undergone peer review. Only 1, nonrandomized study compared the quality of articles published in peer-reviewed and non-peer-reviewed journals. The others measured the effects of variations in the peer-review process or used a before-and-after design to measure the effects of standard peer review on accepted articles. Eighteen studies measured the quality of reviewers' reports under different conditions such as blinding or after training. One study compared the time and cost of different review processes.

**CONCLUSIONS:** Until we have properly defined the objectives of peer-review, it will remain almost impossible to assess or improve its effectiveness. The research needed to understand the broader effects of peer review poses many methodologic problems and would require the cooperation of many parts of the scientific community.

PMID: 12038912

[Indexed for MEDLINE]



*Learned Publishing* (2001)14 , 257–263

# *Shortcomings of peer review in biomedical journals*

**Elizabeth Wager**
*Sideview*

**Tom Jefferson**
*UK Cochrane Centre*



*ABSTRACT: Peer review is well established across most academic disciplines, and publishers, editors, and researchers devote considerable resources to it. This paper uses examples from biomedical journals to examine its shortcomings. Although mainly anecdotal, the evidence suggests that peer review is sometimes ineffective at identifying important research and even less effective at detecting fraud. Most reviewers identify only the minority of a paper's defects and they may be biased. Peer review plus other editorial processes are associated with improvements in papers between submission and publication, but published papers remain hard to read and a significant proportion contain errors or omissions. While it is hard to quantify the costs, peer review does not seem an efficient use of resources. Research into the outcomes of peer review, the establishment of sound methods for measuring the quality of the process and its outcomes, and comparisons with alternative methods are needed.*

# More – The future of scientific publishing?

- For 10 manuscripts, a closed on-line forum of 100 scientists as a crowd source of review was compared to conventional peer review. In all cases, the crowd response was more than enough to enable a fair editorial decision, and compared to the conventional review, the crowd was much faster (days vs months) and collectively provided more comprehensive feedback.

WORLD VIEW *A personal take on events*

DAVID AUSSERHOFER

## Crowd-based peer review can be good and fast

*Confidential feedback from many interacting reviewers can help editors make better, quicker decisions, explains* **Benjamin List**.

# More – The future of scientific publishing?

- Scientific publishing is already using some AI technologies to: identify new peer reviewers; fight plagiarism; ensure all necessary information is reported correctly; find bad statistics; detect data fabrication; verify author identities; suggest keywords; predict IF.

  In the future, software will be able to complete subject-oriented reviews that will enable a fully automated publishing process – including the decision to publish.